C O P *i* L CAMBRIDGE OCCASIONAL PAPERS *in* LINGUISTICS
Volume **7**, Article 4: 83–134, 2015 | ISSN 2050-5949

THE CONCEPTUAL NETWORK APPROACH TO SEMANTIC TYPOLOGY: INTRODUCTION AND APPLICATION

JAMES N. CHANCHARU
University of Cambridge

ABSTRACT This paper introduces a new approach to the comparative study of linguistic meaning, called the conceptual network (CN) approach, and illustrates how this approach works by applying it to an actual data set. This newly proposed approach is especially designed to overcome the disadvantages inherent in the two previous approaches to semantic typology, i.e. the semantic map (SM) model (e.g. Haspelmath (1997)), and the multidimensional scaling (MDS) model (e.g. Croft and Poole (2008)). Specifically, unlike SM, the CN approach does not rely on the perfect fit principle and is computationally tractable, hence the ability to cope with a much larger data set. Furthermore, unlike MDS, the CN approach offers a more accurate way of representing geometrically the conceptual proximities between the functions. Moreover, language universals, in the form of implicational hierarchy, can be represented in a CN model in in a way that the principles behind the model remain consistent. When applied to Haspelmath's indefinite pronoun data, the approach can produce satisfactory results, and some interesting tendencies previously neglected become discernible. However, wider applications of the approach are needed in order to test its validity as a tool for semantic typology, and potentially for other fields of linguistic study.

1 INTRODUCTION

This paper introduces a new approach to the comparative study of linguistic meaning, called the conceptual network (CN) approach, and illustrates how this approach works by applying it to an actual data set of indefinite pronouns. In fact, in the semantic typological literature there already exist two main approaches to such a study: the semantic map (SM) model (e.g. Haspelmath (1997)), and the multidimensional scaling (MDS) model (e.g. Croft and Poole (2008)). These two previous approaches, however, suffer from a number of disadvantages. Most notably, the SM model is a probability-free model that operates in a discrete fashion and does not take in account probabilistic information. As such, the model is less than ideal when implicational universals,

©2015 Chancharu
This is an open-access article distributed by the Department of Theoretical & Applied Linguistics, University of Cambridge under the terms of a Creative Commons Non-Commercial License (creativecommons.org/licenses/by-nc/3.0).

in the form of statistical tendencies, are to be accommodated. On the other hand, the MDS is constraint-free, i.e. it represents non-discrete information in the form of points with different degrees of geometric distances between them, representing different degrees of conceptual proximities. However, by design, the model suffers from the problem of inaccuracy and does not offer a principled way to induce the universals and variations from its non-discrete presentation. The rationale for the introduction of the CN approach, therefore, is to overcome the disadvantages inherent in the previous approaches. Specifically, by incorporating probabilistic information, this new approach aims to formulate a set of predictions about the encoding behavior of the type of constructions investigated. In order to introduce this approach, relevant discussion and explanations are presented in the following sections. Section 2 begins with a brief description of semantic typology as a field of study, followed by the introduction of the SM approach and its disadvantages. Then, Section 3 illustrates how the MDS approach operates. Some of the drawbacks of the MDS approach are presented in Section 4. Section 5 introduces the CN model by describing its components and principles. Furthermore, a case study of the polyfunctionality of indefinite pronouns is described in Section 6. Finally, Section 7 concludes the paper by summarizing the main points made and suggesting further research into the application of the newly proposed model.

2 Polyfunctionality, semantic typology, and the semantic map approach

Let us first consider some key terms and notions relevant to the analysis of this paper. The understanding of the term 'function' as employed in this study presupposes the notion of 'construction,' which can be simply defined as a 'form–function pairing'. In semiotic terms, the construction corresponds to the notion of 'sign', i.e. a symbolic unit whose ability to transfer information consists in the relation between the 'signifier,' i.e. the form of the sign, and the 'signified,' i.e. its function. It should be noted that, like that of the sign, the ontological conception of the construction is not actually physical, but involves a considerable degree of abstractness. That is, by form what is meant here is the mental representation of the phonological specification or morphosyntactic schema of the construction, rather than the actual sound or utterance produced. Moreover, the function refers to the conceptual representation associated with the construction, rather than the actual entity in the physical world. However, the range of linguistic phenomena that can be regarded as constructions potentially goes beyond Saussure's (1916) original conception of the sign, which is generally restricted to the 'instantiative' level. In fact, a construction may also have a relatively 'schematic' representation. That is, it may formally represent a grammatical configuration rather than a specific phonological shape. For example, while the English noun *tree* is a lexical-level construction highly specified in terms of phonology, the English ditransitive construction [S V IO DO] is schematic in that only morphosyntactic specifications (e.g. word order, word classes, and grammatical relations) are posited. Indeed, a wide range of lexical items that meet the specifications can fill in the slots and express the constructional content X CAUSES Y TO RECEIVE Z. It should be noted that the notion of function as adopted in this study is highly compatible with many linguistic theories that place central importance on the exposition of constructions and their relations in grammatical analysis, such as Goldberg (1995; 2006), Langacker (1987; 1991; 2008), and Croft (2001).

Furthermore, more often than not a construction may exhibit 'polyfunctionality', i.e. the phenomenon whereby the same linguistic form is associated with multiple, related functions. In other words, the form–function relationship of a polyfunctional construction is one-to-many, with its multiple functions forming a network of relations. It should be noted that the notion of function as employed in this study goes beyond the lexical semantic level. That is to say, the function of a construction may involve not only the conceptualization of an entity or process (i.e. lexical meaning), but also the relation between different entities and/or processes or between the event conceptualized and the conceptualizer (i.e. grammatical function). Also, functional multiplicity may be attributed to the communicative interaction in a communicative event between the speaker and the listener (i.e. pragmatic function), or the interaction between different, connected communicative events (i.e. discursive function).[1] This multilevel approach to the identification of functional multiplicity is compatible with a number of cross-linguistic studies that compare the functional distributions of constructions of a specific class, such as Kemmer (1993), Haspelmath (1997), and van der Auwera and Plungian (1998). Moreover, for a construction to be considered polyfunctional, its multiple functions must exhibit a considerable degree of relatedness such that they form a network of 'polysemy.'[2] That is to say, those multiple functions must be related in

1 An example of a lexical item that exhibits polyfunctionality is the Thai verb *ʔaw*, which basically means 'to take (something) into one's possession.' Besides its lexical meaning, the form can also serve a grammatical function as a subordinating conjunction meaning 'so much that' and as a preposition meaning 'not until.' Moreover, this same form can be used as a pragmatic marker, either in isolation or preceding another utterance, to express the speaker's negative judgment regarding the listener's surprising behavior, and also as a discourse marker that prompts the listener to start responding to the speaker's directive act, e.g. by answering the question asked.

2 Closely related to but different significantly from polysemy is 'homonymy,' a phenomenon in which two or more constructions happen to share the same form but have unrelated sets of functions. This lack of relatedness is probably evidence that the constructions have different historical sources, or that, if they actually come from the same historical source,

such a way that conceptual association between them is possible and historical pathways between them are reconstructible. Such conceptual associations can be attributed to the cognitive ability of categorization, which operates by such principles as family resemblance and prototype effects. Indeed, this approach to the identification of functional relatedness is compatible with a number of studies in the cognitive semantic tradition that seek to explain the polysemy networks of specific constructions, such as Lakoff (1987), Tyler and Evans (2003), and Evans (2004).

Indeed, the exploration of how meaning is encoded in different languages is by no means a new enterprise. Anthropological studies of kinship terms by Lounsbury (1964), and of color terms by Berlin and Kay (1969), are examples of early attempts to address the question of how the same reality can be carved up in different ways by different linguistic systems, and how such different ways of carving up the same reality probably determine or influence human perception and conceptualization of the world. However, the current tradition of comparative research into meaning, under the name semantic typology, generally aims to formulate universals and variations of linguistic meaning and meaning-encoding strategies as being reflective of the possibilities and limitations of human cognitive functioning. In addition, assuming that human language is in essence a type of sign system, though undoubtedly the most elaborate and complex of such systems, the typological approach to linguistic semantics can be viewed as part of semiotic typology. Specifically, according to Evans (2010: 504), semantic typology can be understood from a semiotic perspective

the constructions have diverged in such a way that they probably have separate mental representations at the synchronic level. An example of the latter case is English *spring*. As a noun, it has quite a few meanings, e.g. 'the season between winter and summer' and 'a twisted piece of metal.' Though historically related, at the synchronic level these two meanings are so conceptually distant that making a mental association between them can be difficult. Therefore, at least two different constructions, rather than a single polyfunctional construction, are represented by the form *spring*. Furthermore, as the relatedness among functions is usually not easy to determine, distinguishing between cases of polysemy and homonymy can be problematic. Specifically, the functional multiplicity of a certain form can be a product of historical accident: that is, the form actually represents two or more constructions derived from different etymological sources. Nevertheless, the multiple meanings of the form may synchronically evoke a conceptual link. An example is Thai *phan*, which can mean 'twine, tangle' or 'species, lineage.' These two meanings, though sharing the same phonological form, are traditionally realized by different spellings, reflecting their historical independence. While the former is a native lexeme, the latter is a Sanskrit loan (*bandha*, cf. English *bond*). However, to some speakers, the two meanings may be linked, probably via conceptual metaphor, i.e. if two things are twined together, they are connected and perhaps related in one way or another. For those speakers who make a conceptual association between the two meanings, it is possible that they merge the two different constructions into one. Evidence for this merging is to be found in the fact that some Thai speakers confuse the two spellings, i.e. they use the orthographic form that originally means 'species, lineage' in a context where the meaning is 'twine, tangle', and vice versa.

as 'the systematic cross-linguistic study of how languages express meaning by ways of signs', which can be distinguished into lexical, grammatical (i.e. morphological and syntactic), and prosodic signs. In other words, a semantic typological study typically deals with a cross-linguistic comparison of meaning encoded by lexical words, morphological items, syntactic patterns, and/or prosodic features. As such, a semantic typological study often incorporates insights from lexical, morphosyntactic, and prosodic typology, although its focus is the comparison and categorization of meanings encoded by lexical, morphosyntactic, and/or prosodic strategies across languages, rather than the comparison and categorization of those strategies themselves.

The past few decades have witnessed the development of a typological approach that is particularly useful for the task of comparing the semantics of linguistic signs across languages. Instigated by Anderson (1974), it is generally referred to as the SM model, and has been applied in several influential typological works, including Kemmer (1993) on passive and middle voices, Haspelmath (1997) on indefinite pronouns, and van der Auwera and Plungian (1998) on modality. It should be noted that semantic mapping could be characterized as a distinct approach in a number of theoretical and methodological aspects. Specifically, it hinges upon a broadly functional linguistic assumption that language can best be described as a communicative tool that is used to perform particular functions at different levels of meaning representation (e.g. discursive, pragmatic, semantic), and that language can best be understood by identifying those functions and studying how they can be represented and performed by linguistic means. In addition, the SM approach is based on a linguistic typological assumption that universality in human language can be uncovered by comparing the strategies that different languages employ in order to fulfill communicative functions. Lastly, the SM approach is also aligned with cognitive linguistics in the sense that the geometric representation produced by means of semantic mapping is considered to reflect the mental representation of linguistic knowledge, especially in relation to polysemy and polyfunctionality, and the cognitive processes concerned.

Most studies conducted using the SM approach are similar in that their findings are framed by the basic components of the SM model, which are 'conceptual space' (also known as conceptual domain or semantic domain) and 'constructional matrix' (also known as lexical/grammatical matrix or language-specific/construction-specific SM). First, a conceptual space is the graphic representation of a particular functional domain. As such, it is conceived of as containing a number of related functions, be they semantic or pragmatic or discursive, that can be encoded by a class of comparable constructions. Each of the functions contained in a particular conceptual space is represented as a functional node holding a position relative to each other. Furthermore, any two functional nodes may be represented as contiguous or connected following Croft's (2001: 96) 'semantic map connectivity hypothesis', which states that 'any relevant language-specific and/or construction-specific category should map onto a connected region in conceptual space'. A connection established between any two functions by virtue of their being encoded by the same linguistic form without implicating the encoding of any other function may be represented by either spatial contiguity or a linking line. Where there is historical evidence regarding the direction of change that relates the functions diachronically a directional arrow can be added to the line. Then, a constructional matrix is drawn to represent the functional range of a particular linguistic construction, be it lexical or grammatical, in the form of a subgraph that geometrically contains the connected functional nodes associated with that construction. It can be said that the final result is a model that translates an implicational hierarchy of related linguistic functions into a geometric representation.

An SM model, according to de Haan (2010), can be constructed in two different directions. In a top-down approach, a particular functional domain is assumed to form a conceptual space that contains a predetermined set of related functions. These functions are predetermined in the sense that they are assumed to be unique and distinguished from each other by a particular set of parameters. In other words, the functional distinction is made in an a priori fashion by taking into account semantic/pragmatic/discursive features that are relevant to the functional domain investigated. To illustrate, in van der Auwera and Plungian's (1998) SM model of modality, eight modal functions are distinguished using four parameters. The first is the strength of modal force (i.e. possibility vs. necessity). The second parameter is whether a modal function is epistemic or not. Third, non-epistemic modality can be either participant-internal or participant-external. The last parameter is whether participant-modality is deontic or not. Then, the matrix of each of the comparable constructions from different languages, presumably representing its functional range, is drawn onto the space. For example, the functional nodes of participant-internal possibility, participant-external, non-deontic and deontic possibility, and epistemic possibility are covered by the subgraph of English *may*, whereas Dutch *mogen* is restricted only to deontic possibility, as illustrated in Figure 1. It should be noted that in the top-down approach to semantic mapping, what determines the distinction between functions is not necessarily based on empirical data. Therefore, it is possible that some functional nodes are treated as distinct even when there is no pair of constructional matrices that differ with respect to them. This is because their distinctiveness has been predetermined by the researcher in postulating the relevant parameters.

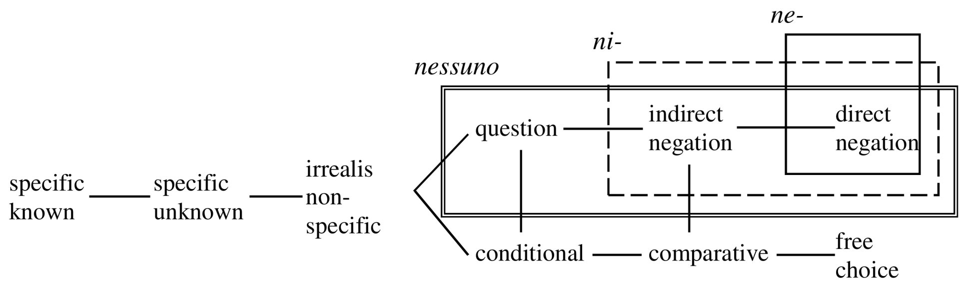

Figure 1 Van der Auwera & Plungian (1998) SM model of modality, with the constructional matrices of English *may* (horizontal striped) and Dutch *mogen* (vertical striped)

On the other hand, an SM model constructed in a bottom-up direction starts from the constructional-matrix level. That is, the functional ranges of a particular set of comparable constructions from different languages are compared, and a functional distinction is made if there is at least one pair of languages that differ with respect to this function. For example, in Haspelmath's (1997) SM of indefinite pronouns, Italian *nessuno* and the Romanian *ni*-series are functionally different in that the former can be used in the interrogative and direct and indirect negative constructions, but the latter can be used only in the direct and indirect negative constructions. Therefore, the question function can be identified as a distinct node in the space. It should be noted, however, that, at this point of the analysis, the direct and indirect negation functions cannot be methodically distinguished from each other, and would be treated as belonging to the same functional node, probably as simply negation. It is only when certain constructions, such as those Latvian indefinite pronouns that have the negative prefix *ne*-, are taken into consideration that direct and direction negation can be further distinguished, as the Latvian *ne*-series can be used only in the direct negative context, as shown in Figure 2. In this respect, the conceptual space can be said to emerge consequently from this process of functional identification by means of comparison and contrast. Nevertheless, there is no reason to presume that the top-down and bottom-up approaches to semantic mapping are incompatible with each other. Indeed, Zwarts (2010) has even suggested that, to overcome certain specific disadvantages inherent in each of the two approaches, they be applied complementarily in constructing an SM model.

Still, there are concerns that the SM approach might, in some cases, not be an effective tool of typological analysis. Specifically, Croft and Poole (2008)

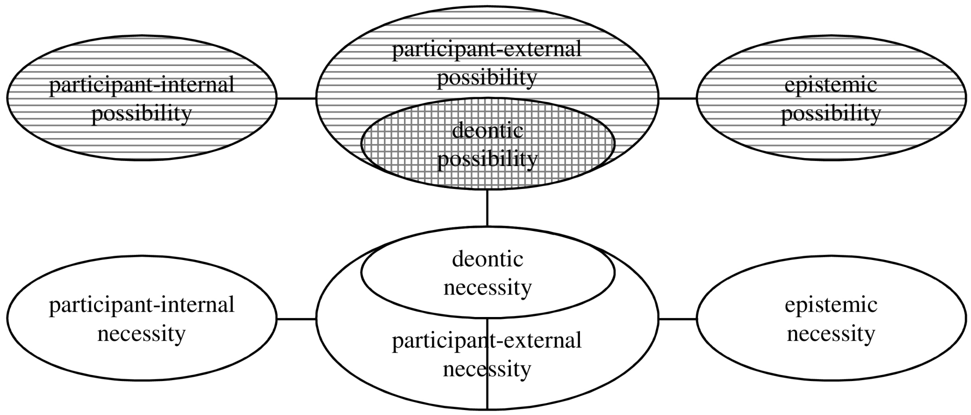

Figure 2 Haspelmath's (1997) SM of indefinite pronouns, with the constructional matrices of Italian *nessuno* Romanian *ni-*, and Latvian *ne-*

argue that the SM model lacks the capacity to deal with large cross-linguistic datasets that exhibit a high degree of variation. According to Croft and Poole, three specific problems can be identified: first, constructing an SM model does not involve any systematic mathematical method or device, but is traditionally carried out in a manual fashion. Therefore, it is difficult for a SM model to accommodate a large number of data points. Indeed, even when a relatively small set of data points is involved, a model that is constructed manually is prone to error. Croft and Poole reported one possible error discernible in Haspelmath's model, in which the functional connection between the irrealis non-specific and conditional functions is established unnecessarily. The second problem with the SM approach is that it is entirely based on the 'perfect fit' principle. That is, every set of functions that are encoded by the same linguistic construction must be represented graphically as connected, either directly or indirectly. Such a perfect fit model, however, is not very useful in view of probabilistic computation, and in fact 'may be theoretically less informative than a high but not perfect fit' (2008). Last, the spatial representation of an SM model does not lend itself to mathematical interpretation. This is because the configuration of the components of the model is achieved manually without any statistical method or device. To illustrate this by reference to the case we have just been discussing, although the connection between the functional nodes of a model is claimed to reflect the mental representation of the respective functions, the actual distance between those nodes has no mathematical significance. In relation to this, it is questionable whether 'the semantic map model can be automated in a computationally tractable algorithm' (2008).

3 The multidimensional scaling approach

As an alternative, Croft and Poole (2008) proposed a new model, called the MDS model. This new model makes use of the Optimal Classification nonparametric unfolding algorithm, which is a multivariate statistical technique developed by the second co-author (Poole, 2000). When applied to semantic typological analysis, this technique is claimed to be advantageous in that it can 'preserve the theoretical insights of the semantic map model without the attendant problems', as identified above (2008). One of the basic ideas of this technique is that the (dis-)similarities between items of a particular set are represented as distances between points that represent the items in a Euclidean space, with more similar items represented as points closer to each other. These points, according to Croft and Poole, 'form a spatial model that summarizes the similarities/dissimilarities data' (2008). In semantic typological terms, items determined to be either similar or dissimilar to each other are comparable to linguistic functions determined to be encoded either by the same construction or by different constructions. Also, the points representing the items are comparable to the functional nodes in the conventional SM model. However, in the MDS model, the actual distances between the points are claimed to be significant for, and even central to, the task of inferring linguistic universals. In addition, the representation of a language-specific, construction-specific functional range takes the form of a geometric region bounded by a line that cuts across the conceptual space.

As an illustrative application of MDS, Croft and Poole (2008) offer a re-examination of Haspelmath's (1997) data. The functional distributions of 139 pronouns from 40 languages across nine functions are analyzed using the Optimal Classification nonparametric unfolding algorithm. The process of analysis essentially involves the spatial configuration of the ideal function points along the cutting lines representing the functional distributions of the constructions examined. Geometrically, a cutting line is a line that is drawn onto the space in such a way that it passes through the midpoint of the distribution of the data points relevant to the construction represented (equivalent to the median in statistical terms), and that it is perpendicular to the normal vector line (equivalent to the normal distribution curve in statistical terms). Ideally, a cutting line divides the data points into Y and N regions in an accurate fashion. On one side of a cutting line are function points that have the Y value, i.e. that are encoded by the construction, while on the other are those with the N value, i.e. that are not encoded by the construction. For example, the line that represents Romanian pronouns with the prefix *ni-* cuts the geometric space into two parts: the part that contains all the functions encoded by the pronoun, i.e. direct and indirect negation, and the part that contains all the functions not encoded by it – as illustrated in Figure 3. Next, when more constructions are included in the computation, the possible positions of the data points are readjusted in order to accommodate both the old and new data. For instance, when the cutting line that represents the Romanian *vre-un* indefinite pronouns is drawn, it must delimit a region in which the possible function points encoded by the pronouns, i.e. conditional, question, and indirect negation, are located. In this way, the function encoded by both the *ni-* and *vre-un* series, i.e. indirect negation, needs to be represented in a region defined by the intersection between the *ni-* and *vre-un* cutting lines. When more constructions are taken into account, e.g. the Romanian indefinite pronoun series with *-va* ad *ori-*, as shown in the figure, the space is further segmented into regions. It should be noted that the conceptual space in which the process of positional adjustment operates is a multidimensional Euclidean one, i.e. it involves the positioning of the lines and points in many different spatial dimensions.

Furthermore, when indefinite pronouns from languages other than Romanian are included in the computation, more cutting lines are drawn onto the space. These cutting lines conceptually represent the boundaries of the functional ranges of the constructions investigated, and their intersections 'define regions (called polytopes) within which the ideal point of the function is located' (2008). With more intersecting lines drawn, each of the polytopes tends to get smaller. A smaller polytope means a more limited range of possible sites for the ideal point of the function. In other words, the precision of the positioning of an ideal point depends on the number of the cutting lines. However, more often than not the polytopes represented on the same space can vary greatly in terms of area. That is, some polytopes are so small that the ideal point of a particular function can be determined in a somewhat precise fashion. Conversely, some polytopes are defined by fewer intersecting lines and so are of a considerable surface area. Consequently, the range of possible sites for the ideal point is relatively large – hence a lesser degree of precision. As illustrated in Figure 4, the regions that contain the possible ideal points of the functions are of varying sizes. Most notably, the possible ideal points of the irrealis non-specific and indirect negation functions are contained in regions that are quite large, and so the actual sites 'could be anywhere in the regions' (2008). In other words, the function points shown in the figure are presented at their potential, as opposed to decisive, positions, and in the case of larger regions, like those of the irrealis non-specific and indirect negation functions,

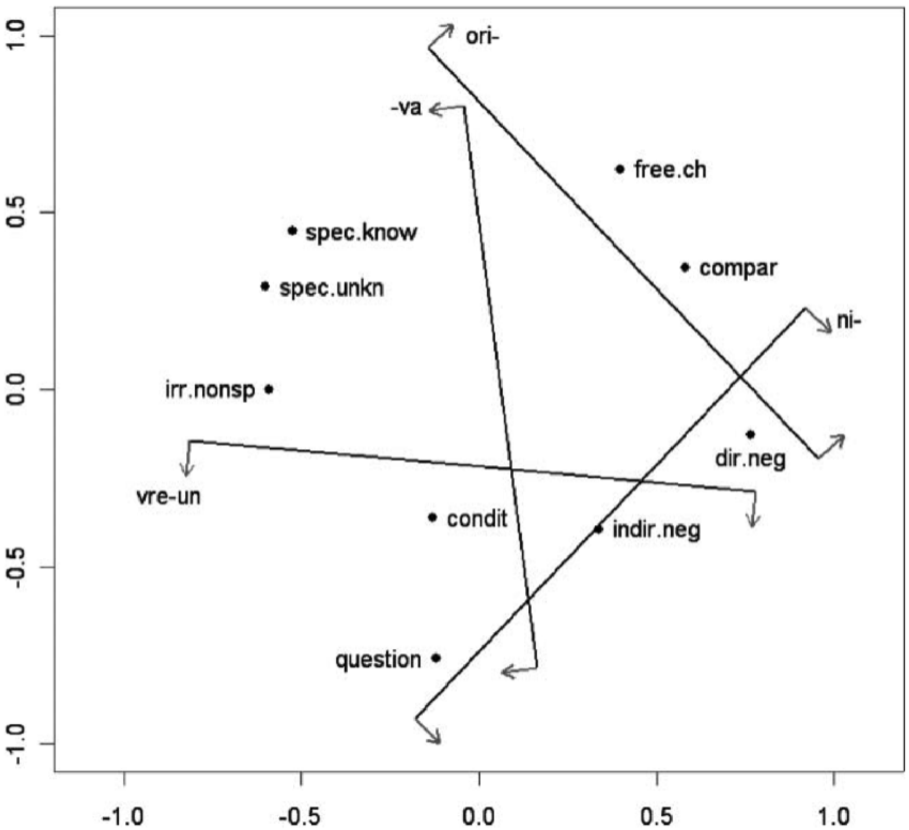

Figure 3 Croft & Poole's (2008) MDS model of indefinite pronouns (based on Haspelmath's (1997) data), with the cutting lines for four classes of Romanian indefinite pronouns

their representative points are positioned in a somewhat random fashion.

Upon the completion of the process of spatial configuration, a measurement of the fitness of the model is performed. According to Croft and Poole (2008), one fitness test that can be applied to a MDS model is correct classification of the data. Specifically, the test of correct classification measures the degree to which 'the cutting lines correctly separate Y and N values' (2008). A case of misclassification, however, occurs when the position of a cutting line fails to correctly represent the division of the data into Y and N regions, especially when the functional distributions of many constructions in the data exhibits a low degree of linear dependence. The correctness of classification is presented in percentage form, with a higher percentage meaning that the model can

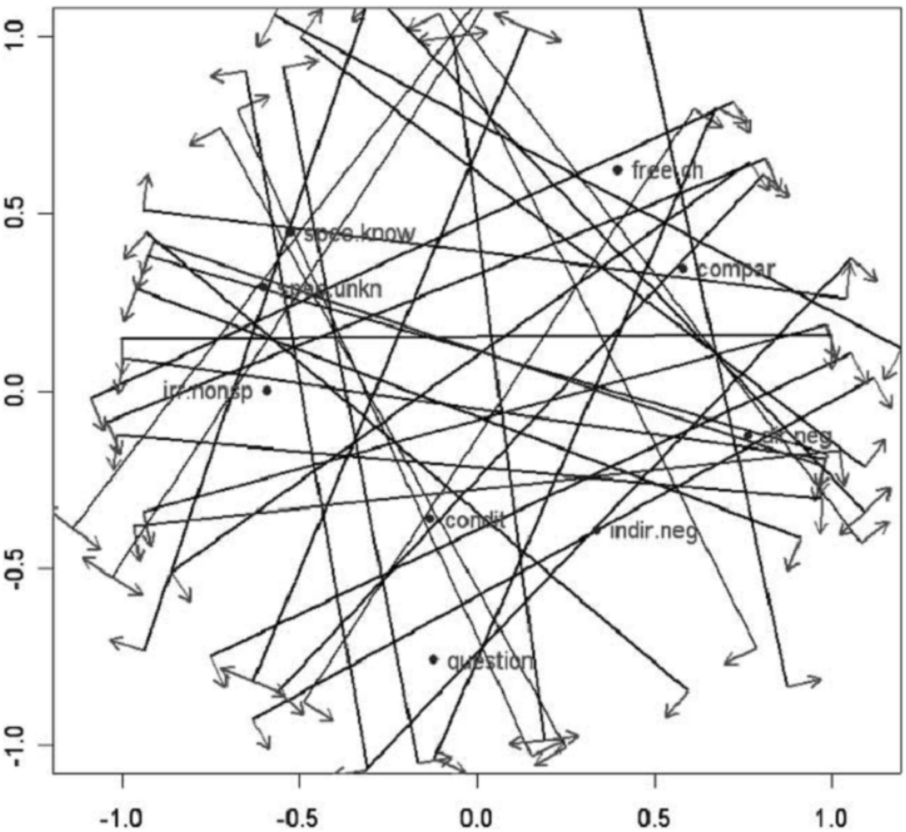

Figure 4 Croft & Poole's (2008) MDS model of indefinite pronouns, with the cutting lines representing all the 139 constructions, many of which are identical

better accommodate the classificatory system imposed by the positioning of the cutting lines and ideal points. Also, the test of correct classification can be applied to determine in how many dimensions the model is best represented. That is, as the process of spatial configuration operates on a multidimensional Euclidean space, a different number of dimensions of representation tends to have a different degree of correctness of classification. Generally, a higher degree of dimensionality guarantees a more correct classification, as an additional dimension means one more way to configure the positions of lines and points to match the Y and N values of each function and construction. However, a model with greater dimensionality is not always preferable because complexity is also added. Ideally, the better model is the simpler one, and so it is possible that the increase in data fit gained by adding an extra dimension is considered too small to justify the loss in terms of added complexity. Moreover, a model with a perfect fit (100% correct classification) is disadvantageous in that it is 'not the usual state of affairs for models of complex human behavior', and therefore 'may be theoretically less informative than a model with a high but not perfect fit' (2008). Therefore, a rule of thumb for the best representation is probably the lowest degree of dimensionality with the highest degree of improvement of fit. To illustrate, Croft and Poole (2008) reported that the conceptual space constructed from Haspelmath's (1997) data may be represented in a one-, two-, or three-dimensional model with degrees of correct classification of 90.8%, 98.1%, and 100%, respectively. In this case, although the three-dimensional model gives a perfect (100% correct) classification, there is only a slight improvement (1.9%) from the two-dimensional model. Therefore, the best model of representation would be the two-dimensional (98.1% correct), offering the highest degree of improvement of 7.3%.

4 Problems with the multidimensional scaling approach

Although Croft and Poole's (2008) MDS approach has a number of advantages over the SM model, as described above, it is still not satisfactory in the descriptive and explanatory domains. First, MDS is descriptively inadequate when applied to relatively small data sets. This inadequacy follows from the fact that the Optimal Classification nonparametric unfolding algorithm developed by Poole is in essence 'an approximation method', the final output of which results from the 'successive approximations of the positions of the cutting lines and the points' (2008). Technically speaking, the position of the possible ideal point for a particular function is approximate in the sense that it could be anywhere in the region that is defined by the intersection of the cutting lines representing the functional ranges of relevant constructions. Then, with additional cutting lines, the region becomes smaller, and the position of the function ideal point becomes less approximate, i.e. more precise. In other words, a higher precision of the positions of the points requires a large amount of data to be included in the computation. Conversely, with a small data set, the positions of the points representing the various functions in the geometric space are not precisely determined, but remain only tentative. The need for a large amount of data in order to produce precise results can be problematic because semantic typological studies in general rely on a relatively small number of data points, e.g. 1,251 in Haspelmath's (1997), compared to the much larger numbers of data points analyzed by studies in other social sciences, e.g. 23,097 in Poole's (2000) study of spatial models of US parliamentary voting, from which the algorithm was originally developed.

Indeed, this problem was admitted by Croft and Poole (2008) themselves when analyzing Haspelmath's (1997) data. The lack of positional precision with a small data set has at least one particular undesirable effect. Specifically, as shown in Figure 5, the indirect negation function point is located closer to the conditional function point (Distance A) than to the question function point (Distance B), despite the fact that actually there is no construction in the data that encodes the indirect negation and conditional functions without encoding the question function. In statistical terms, the indirect negation and question functions have a correlation value of 0.443 – in fact, greater than the correlation value of 0.386 between the indirect negation and conditional functions. In other words, in Croft and Poole's model, the position of the indirect negation function point relative to the positions of the other function points does not lend itself to a conceptual interpretation, as the distances between them are not actually geometrically significant. This problem results from the fact that the relevant data set is so small that too few cutting lines are produced, resulting in a relatively large polytope. Therefore, the range of possible sites for the ideal point of the function is so large that a misinterpretation of the data may arise. That is, the indirect negation function may be wrongly interpreted as having a direct connection with the conditional function. With this lack of positional precision, it is difficult for linguistic universals to be inferred from a MDS model. For example, an implicational hierarchy like 'if function X is conceptually connected to function Y with D distance, then X is also conceptually connected with any other function with a distance less than or equal to D' cannot be postulated, as the relative distances in the model cannot be translated into conceptual proximities in a reliable way.

It should be noted that, according to Croft and Poole, the mathematical algorithm applied in their MDS model 'is designed to maximize correct classification, that is, the accuracy of the categories defined by the cutting lines' (2008). These categories are geometrically represented by the polytopes defined by the intersections of the cutting lines, and each of these polytopes constitutes a range of possible sites for the ideal point that represents a particular linguistic function. In other words, the proposed algorithm primarily involves the spatial configuration of the cutting lines and the polytopes defined by their intersections. The positioning of the ideal points, then, is epiphenomenal in the sense that it is subject to the density of the cutting lines and size of the polytopes. With more cutting lines drawn onto the space, the polytopes tend to be smaller. A smaller polytope means that there is a smaller range of possible sites for the ideal point of a particular function, hence a higher degree of precision. However, it is possible for an MDS model to have a high degree of correctness without necessarily having a high degree of precision. This is because the correctness of an MDS model depends on the accurate

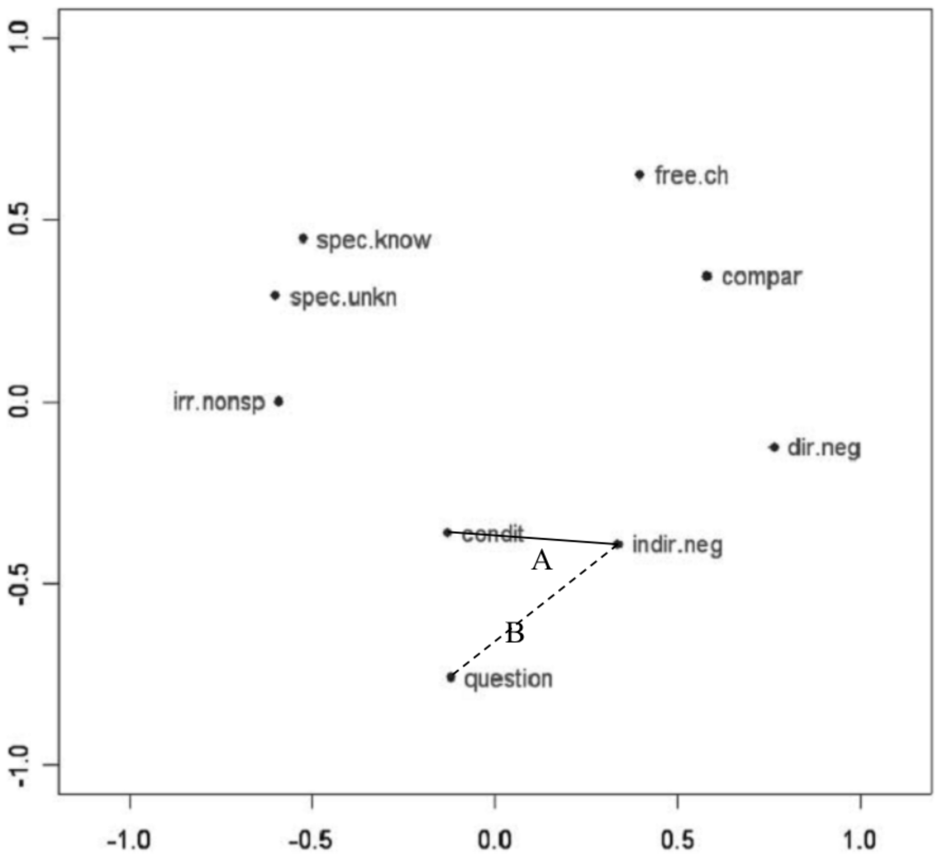

Figure 5 Croft & Poole's (2008) MDS model of indefinite pronouns, featuring Distance A between the indirect negation and conditional function points, and Distance B between the indirect negation and question function points

positioning of the cutting lines and their corresponding polytopes, and not on that of the possible ideal points. An example would be the somewhat random positioning of the possible ideal point of the indirect negation function, hence a potential misinterpretation of the data, as described above, despite the fact that the model as a whole has a nearly perfect fit (98.1% correct classification). This problem of inaccurate positioning of the data points can present a serious drawback for an MDS model that is claimed to represent interpretable distances between the points.

The second problem with the MDS approach is that it lacks explanatory power when it comes to the conceptual connections between the functions.

According to Croft and Poole (2008), in contrast to the discreteness inherent in the SM model, MDS is a non-discrete method. Although I agree that a non-discrete model potentially offers a more informative means of representing data, in the case of semantic typological analysis there need to be constraints on the interpretation of spatial distance. In the traditional SM approach, such constraints exist in the form of connections between functional nodes. Connectivity may take the form of two functions being positioned contiguously to each other or being linked together by a line. As such, the connectivity principle can serve as a basis for a set of predictions: a connection between any two functions means that at the synchronic level there is probably a strong mental association between the functions, and at the diachronic level, a historical path between the functions is possible. On the other hand, the absence of a connection between any two functions means that they are probably not mentally associated (at least directly) and a change in terms of one function being the source and the other being the target is unlikely. Moreover, connectivity can lend itself to the representation of an implicational hierarchy. That is, for any two functions, X and Y, that are encoded by the same construction, all the functions that are represented as intermediately connected between X and Y are also encoded by the same construction as X and Y (1997).

On the other hand, the MDS model does away with discreteness – hence the absence of constraints on connectivity. Instead, conceptual (dis-)similarity 'is modeled in terms of Euclidean distance between points in the representation' (2008). In other words, relations between functions are now treated as relative, as opposed to absolute, with less distance suggesting a closer relationship. Therefore, in a non-discrete MDS model, every connection between any two points in the space is possible in principle, although with differing degrees of probability. A model like this, however, is insufficient, in that the representation of absolute distance alone cannot determine the optimal way to connect any two function points. For example, as shown in Figure 6, it is possible to establish a direct connection between the conditional and direct negation function points (Route A). Alternatively, the two function points can be connected indirectly via the question and indirect negation function points (Route B). Though sharing the same terminal function points, the two routes make different predictions. That is, Route A predicts that a construction can encode the conditional and direct negation functions without encoding the question and indirect negation functions. Route B, on the other hand, predicts that if a construction encodes the conditional and direct negation functions, it must also encode the question and indirect negation functions. Indeed, it is found in the data that all 22 constructions that encode the conditional and direct negation functions also encode the question and indirect negation functions.

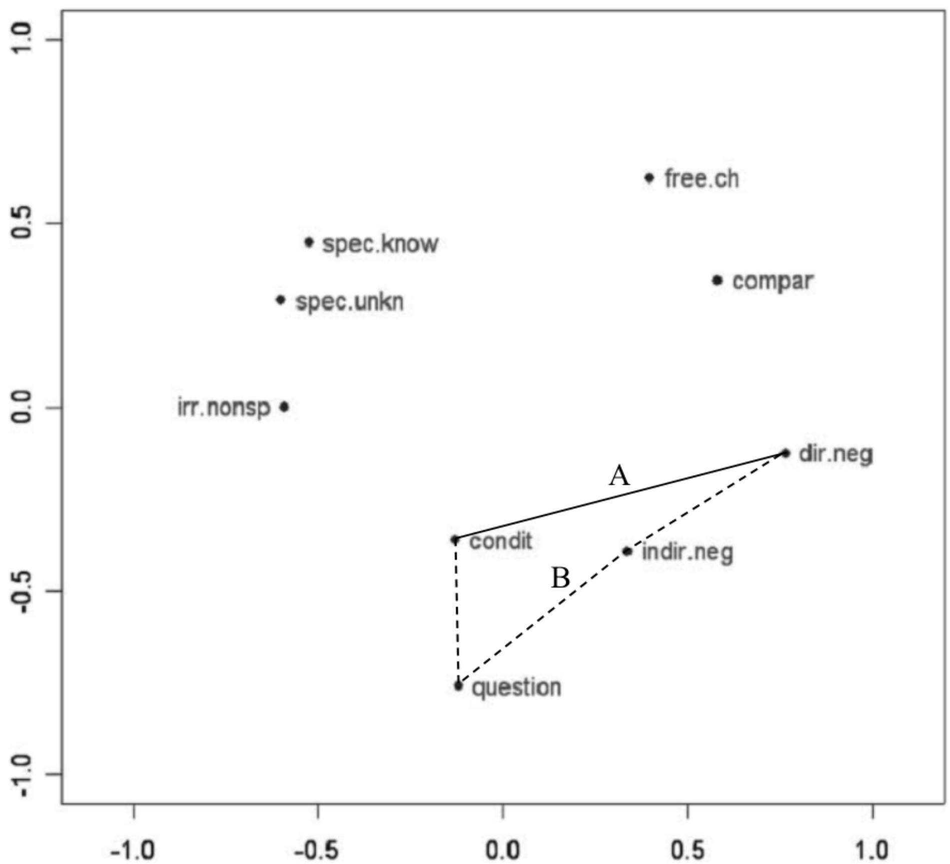

Figure 6 Croft & Poole's (2008) MDS model of indefinite pronouns, featuring Route A directly between the conditional and direct negation function points, and Route B between the conditional and direct negation function points via the indirect negation function point

Furthermore, correlation values between function points show that Route B is preferable. That is, the correlation value between the conditional and direct negation functions (Route A) is significantly lower (i.e. -0.017) compared to the correlation values of 0.857 between the conditional and question functions, of 0.423 between the question and indirect negation functions, and of 0.285 between the indirect negation and direct negation functions (Route B). Therefore, it is clear that, in relation to how they encode indefinite pronouns, presumably languages in general, and indeed all the languages in the data, prefer Route B, and it can be postulated that Route A constitutes an improbable, or even impossible, connection. However, in the MDS approach there is no principled way to explain why this is the case, as practically every connection between any two function points can be established. On the other hand, Haspelmath's (1997) SM model of indefinite pronouns incorporates the information as to which connections are preferred over others in the form of an implication hierarchy. Specifically, it postulates that if a construction encodes the conditional and direct negation functions, it must encode the question and indirect negation functions as well. In this respect, the traditional SM model fares better than the MDS model. In fact, Croft and Poole make an attempt to remedy this disadvantage by suggesting that the connecting lines between the functions of Haspelmath's SM analysis be superimposed on the MDS display, as shown in Figure 7. Nevertheless, the question can be raised as to whether MDS really proves useful for the purpose of semantic typology when eventually the analyst needs to resort to the SM approach for a more precise description of the data.

5 Introducing the CN approach

As a solution to the disadvantages of the SM and MDS approaches to semantic typology, as discussed above, I propose instead a new approach, called the CN approach. In fact, this new approach makes use of a MDS method as well. However, unlike Croft and Poole's (2008) MDS approach, the MDS method used is a different one and forms only the first half of the whole process of data processing. The second half of the process concerns an innovative use of a cartographic technique in generating a network of data points that is assumed to represent conceptual connections between a particular set of grammatical functions, hence the name 'conceptual network.' Basically, the CN approach produces a display that features explicit connections between related functions, by applying a mathematically formalized technique. In other words, CN combines the advantages of both the previous approaches, but nevertheless manages to evade the drawbacks inherent in them. Like SM and MDS, CN is an approach that is designed to deal with empirical data in the form of the functional distribution of a cross-linguistically comparable class of constructions. The data points are in a discrete, binary form, taking the value of either 1 or 0, i.e. whether a given construction can be used to encode a particular function or not. To illustrate, suppose there are twelve constructions (1–12) from different languages that are functionally comparable, i.e. they are more or less associated with the same functional domain. Moreover, five functions (A–E) are identified as belonging to that domain. It should be noted that the identification of the functions can be either a top-down or bottom-up process, or even a two-way process, as described above. The distribution of the twelve constructions across the five functions is shown in Table 1.

Next, a matrix of pairwise comparisons between the functions in relation to

	A	**B**	**C**	**D**	**E**
1	1	1	1	0	0
2	0	0	1	0	1
3	1	1	1	1	0
4	0	1	1	0	1
5	0	1	1	1	1
6	1	1	0	0	0
7	0	0	1	1	1
8	0	1	0	1	0
9	0	0	0	1	1
10	0	0	1	1	1
11	0	1	0	1	0
12	1	0	0	0	0

Table 1 Distribution of Constructions 1-12 across Functions A-E (1=encoded; 0=not encoded)

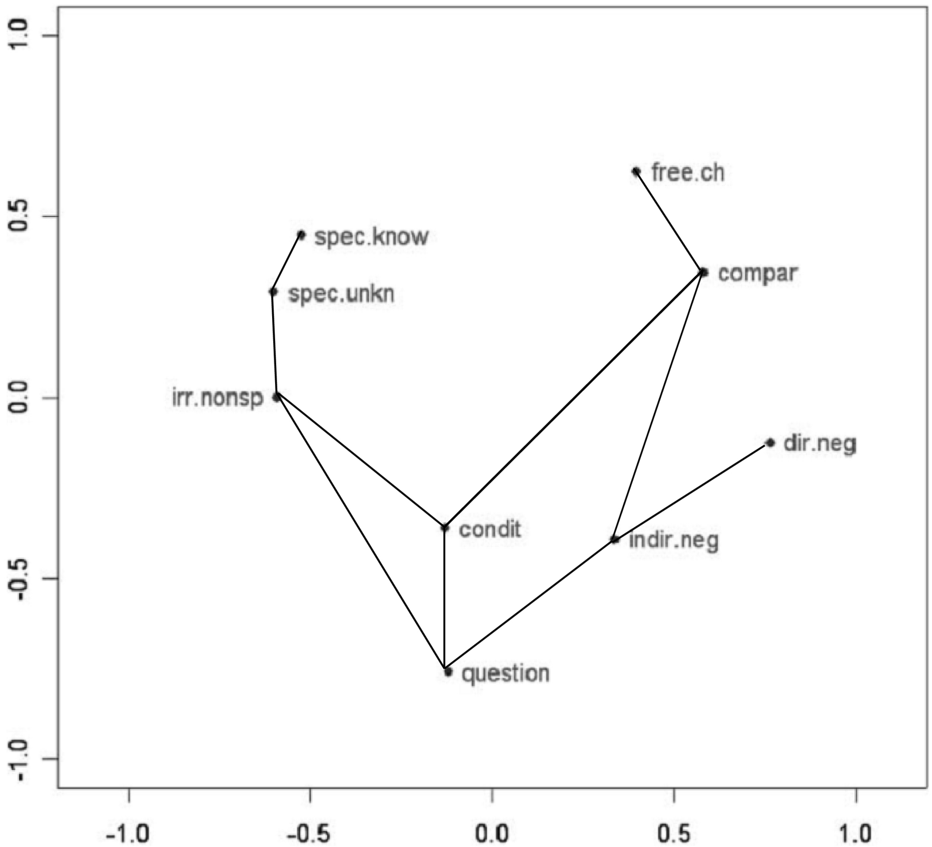

Figure 7 Croft & Poole's (2008) MDS model of indefinite pronouns, with Haspelmath's (1997) SM graph structure superimposed on

each of the constructions is constructed, based on the functional distribution table. Such a matrix contains the values of conceptual proximities between each pair of the functions. In fact, a number of statistical measures of proximity can be employed for this purpose, but the measure adopted here for the sake of illustration is the Pearson correlation coefficient. Table 2 shows the matrix of pairwise comparisons between the functions in relation to each of the constructions in the constructed example, measured by means of the Pearson correlation coefficient. It should be noted that Croft and Poole argued against a method of this kind, that consists in constructing a matrix of pairwise comparisons, claiming that it is time-consuming and may cause loss of information in the case of lopsided data (2008). I argue, however, that there is in fact no time problem: all that is required is a degree of patience on the part

	A	B	C	D	E
A	·				
B	.408	·			
C	-.120	-.029	·		
D	-.478	-.029	-.029	·	
E	-.478	-.371	.657	-.029	·

Table 2 Matrix of pairwise comparisons of Functions A-E, measured by means of the Pearson correlation coefficient

of the researcher. Also, the lopsidedness of data might actually not be a serious problem. According to Croft and Poole, by measuring the proximity value of each pair of functions, ‘information is lost in the process of constructing the pairwise comparisons’. In the linguistic application, lopsided data may take the form of constructions that ‘are used for very few or almost all functions’ in a given functional domain. In such cases, ‘the dissimilarity values will be all very close’, making the conceptual model less informative than it should be (2008).

In semantic typological studies in general, constructions rarely encode very few or almost all functions. In fact, most constructions fall somewhere between the two extremes, i.e. are used for a moderate number of functions in a given domain. For example, of all the 140 constructions in Haspelmath’s (1997) data of indefinite pronouns, 30 constructions encode only one function, and only one construction encodes seven functions, out of the total of eight. Altogether, these extreme cases constitute only about 22% of the total number of constructions – indeed, by no means the majority. Furthermore, I doubt if many other studies of a similar nature would suffer from the problem of lopsided data. This is because if most constructions investigated were to encode very few functions in a given domain, it would be doubtful if these functions really constituted a coherent, unified domain. Moreover, if the majority of constructions investigated were to encode almost all functions in a given domain, it would be doubtful if these functions were actually worth distinguishing. In other words, it is somewhat counter-intuitive in the first place to construct any systematic and meaningful model of semantic typology based on a large amount of lopsided data. Therefore, there is probably no valid reason not to apply the technique of pairwise comparison to the measurement of conceptual proximities between functions.

Actually, the application of the kind of proximity measure technique as

	A	B	C	D	E
A	.000	-	-	-	-
B	.528	.000	-	-	-
C	1.013	.929	.000	-	-
D	1.342	.929	.929	.000	-
E	1.342	1.244	.299	.929	.000

Table 3 Matrix of transformed proximities between Functions A-E (matrix conditional, interval transformation)

described above provides a good solution to the descriptive problem inherent in the MDS approach, i.e. the lack of precision in the positioning of function points in the conceptual space. This is because, in an MDS model, function points are positioned through a series of successive partitionings of the space into two parts by the cutting line representing each functional range. As a result, with a small set of data, the points will be positioned in a somewhat random fashion in vaguely defined polytopes. By contrast, the CN approach relies on a different technique that involves performing proximity measurement of the functions and then using these proximities to determine the optimal coordinates of the function points in the conceptual space. In this way the conceptual distances between the functions can be precisely determined even if the data set is small. It should be noted that the 10 proximity values in the lower-triangular matrix in Table 2 represent conceptual distances between the functions measured by means of the Pearson correlation coefficient. Also, this is a similarity, as opposed to dissimilarity, matrix. That is, the higher the number, the more conceptually similar the functions are to each other. These values, however, are raw input proximities that need to be transformed so that they can be used to determine the conceptual distances. The process of transformation begins with converting similarities into dissimilarities (by multiplying the numbers by -1). Then, the smallest of the dissimilarities is set equal to 0. Finally, the values are normalized such that the weighted squared values equal the sum of the weights. Table 3 shows the transformed proximities between the functions of our constructed example, based on the correlation values measured between the functions.

Next, a multidimensional scaling program called PROXSCAL (Proximity Scaling), available as an add-on in the statistical analysis software SPSS, is applied to the matrix of normalized transformed proximities in order to construct a Euclidean, least-squares representation of the functions, in the

	X	Y
A	-.730	.234
B	-.502	-.212
C	.294	.296
D	.390	-.612
E	.608	.294

Table 4 Final coordinates of Functions A-E in a two-dimensional space

form of a conceptual space populated by function points. The positions of the function points are configured such that the distances between them fit the transformed proximities as closely as possible. This configuration is performed through a series of iterations. At the end of each iteration there is an evaluation of the fitness, which consists in measuring how well the geometric distances generated fit the transformed conceptual proximities. A fitness evaluation can be conducted using various types of measures. For the purpose of non-metric multidimensional scaling analysis, the stress test is particularly relevant. Stress can be described as the normalized, least-squares index of the fit between the distances and the proximities. In other words, a higher level of stress means a greater degree of mismatch, and a lower level of stress means a better fit. After a number of iterations, the process of configuration terminates if the current stress value is smaller than or equal to a given minimum stress value, or if the difference between two consecutive stress values is smaller than or equal to a particular threshold of improvement. In the case of our constructed example, five iterations are completed before the degree of improvement becomes less than the convergence criterion set, hence the termination of the process of spatial configuration. Table 4 shows the final coordinates of the five function points, the distances between which are configured to reflect the transformed proximity values as closely as possible. Finally, based on these final coordinates, the function points are plotted onto the conceptual space. Figure 8 shows a two-dimensional display of this conceptual space.

Dimensionality is another concern as regards the representation of a multidimensional scaling model. To determine in how many dimensions the conceptual model can be optimally represented, the stress levels can be compared. An elbow graph structure like the one shown in Figure 9 is typical of a scree plot of stress analysis, representing a sharp fall in stress (i.e. a sharp rise in fitness) at a certain level of dimensionality, followed by a gradual decline. In this case, although at three dimensions a zero stress value (.000) is achieved, a

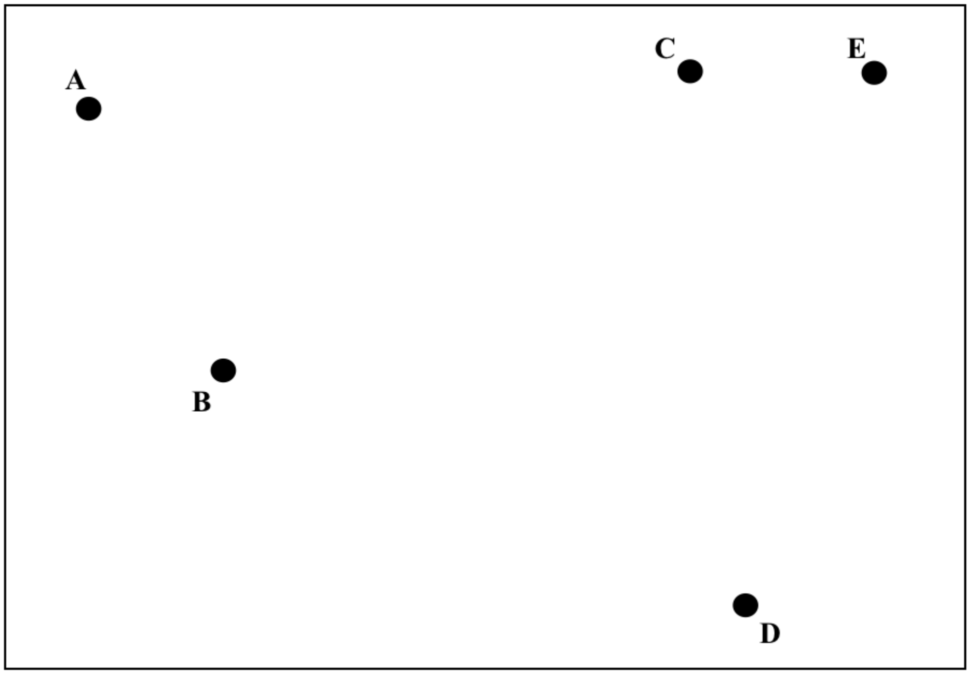

Figure 8 Points A-E plotted onto a two-dimensional space

two-dimensional model is probably an optimal choice. This is because at two dimensions the stress level is also quite low, and the greatest improvement in fitness (i.e. the greatest stress reduction) of .103 is achieved, compared to the improvement rate of .001 at three dimensions. In other words, an additional dimension does not seem to make any significant improvement in terms of goodness-of-fit. It should be noted that although the two-dimensional model does not achieve a perfect fit with the input data, at the stress value of .001 virtually all the conceptual proximities are successfully translated into final distances between the points such that a spatial analysis can be performed using either the former or the latter as input data. Moreover, a model of lower dimensionality is generally simpler, and should thus be easier for the analyst to understand and so should help shorten the data processing time.

Let us now resume the explanation of how to construct a CN model. The second half of the process involves the use of a cartographic technique. For this purpose, I used the geographic information systems application QGIS. It should be noted that with the use of such an application, a number of components can be drawn onto the space such that the model becomes conceptually informative. Specially, four types of components are relevant to the CN model: convex

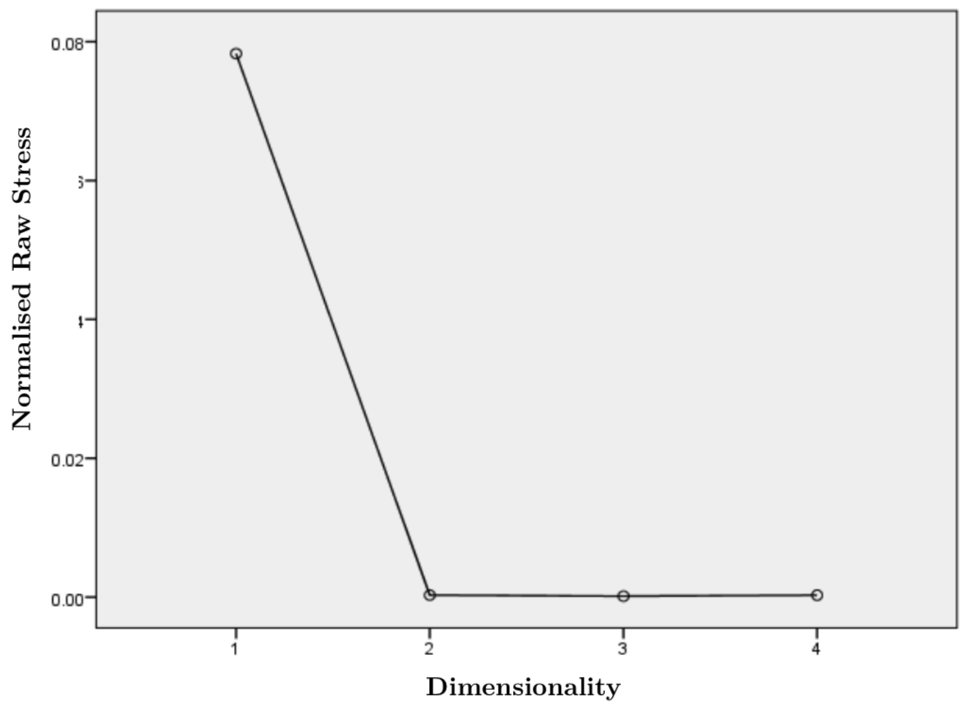

Figure 9 Levels of normalized raw stress by the number of dimensions

regions, prototype points, possible connections, and probable connections. These components are assumed to convey conceptual information based on the corresponding principles of convexity, prototypicality, possibilisticconnectivity, and probabilistic connectivity, respectively. First, the final coordinates produced by the MDS program PROXSCAL are used to create a geometric space, with specific points representing the central properties/examples of the functions examined. Then, the space is segmented into a number of regions by means of a geometrical method called Voronoi tessellation. Each region is essentially a geometric polygon that is defined by the intersection of lines consisting of points equidistant between any two nearest function points. As a result, for each prototype point P, there is a corresponding region that contains every point whose distance to P is less than or equal to its distance to any other prototype point. By means of the Voronoi tessellation technique, a number of connected convex regions are generated around their corresponding prototype points. Indeed, the convexity of the structure of a concept seems to be of both geometric and conceptual importance. Geometrically, a region R is said to be convex if, for every pair of points X and Y within R, every point on the straight line segment that joins X and Y is also within R. See Figure 10 for an illustration of convex (a) and non-convex (b) regions. Figure 11 shows the five prototype points of our constructed example, with their corresponding convex regions, generated by means of Voronoi tessellation.

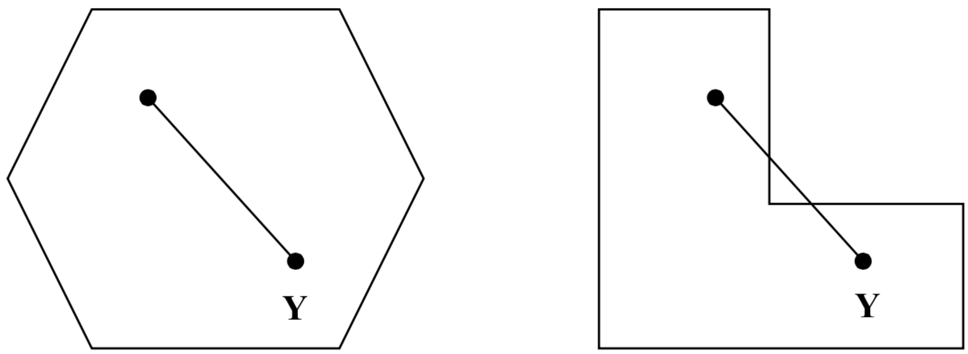

Figure 10 Regions that are convex (left) and non-convex (right)

It is possible that the representation of the prototype of each function as a point contained in a region reflects the cognitive linguistic conception of meaning. Specifically, according to Gärdenfors (2014), a region that is constructed by means of Voronoi tessellation, as described above, can be considered to be a category of linguistic function, with any point contained in the region representing a member of that category. Generally, the various members of a category are by no means of equal status, but differ from one another in terms of prototypicality. Prototypical members share more of the central properties of the category, whereas other members that are peripheral share fewer of these central properties, or have only the minor properties. The conceptual structure of a category, then, asymmetrically centers on its more prototypical membership. In geometrical terms, a Voronoi region centers on the central point that is considered to represent the most prototypical instance of the relevant functional category. According to Gärdenfors, the convex structure of a certain conceptual category could facilitate the learnability of the category and ensure the effectiveness of communication (p. 26). Specifically, Gärdenfors reported on a number of research findings in which the learning of a certain concept becomes more accurate and the communication of a certain concept becomes more effective once the convex representation of the concept is assumed.

Moreover, it is plausible that the segmentation of a conceptual space as a whole into connected convex regions that center on prototype points makes

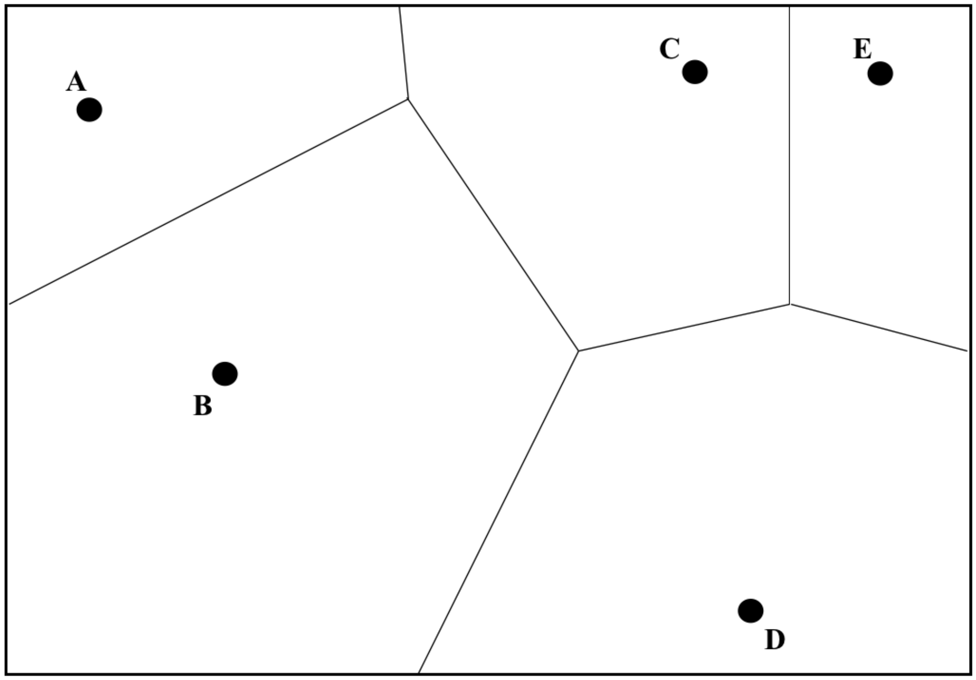

Figure 11 Prototype points A-E and their corresponding convex regions, generated by means of Voronoi tessellation

the process of concept learning and processing more cognitively economical. According to Gärdenfors (2014), it is unlikely that the categorization of each point in a space, i.e. each instance/facet of meaning in a functional domain, is to be memorized, as 'this would put absurd demands on human memory'. On the other hand, 'if the partitioning of a space into categories is based on a Voronoi tessellation, only the relative positions of the prototypes need to be remembered', hence involving much less cognitive load. This is because once the positions of the prototypes are recalled, 'the rest of a categorization can be computed by using the metric of the space' (pp. 27–28). For this reason, the categorization of a functional domain by means of tessellation mechanism can present a more cognitively realistic way to represent the domain spatially. Specifically, this mode of presentation falls in the middle between the extremely discrete case like the SM model, on the one hand, and the extremely non-discrete case like the MDS model, on the other.

Indeed, this adaptation of the Voronoi tessellation technique to the study of semantic typology is innovative. Although Gärdenfors (2000; 2014) already argued that the structure of a concept takes the form of a convex object, and that Voronoi tessellation can be used to segment a conceptual space into convex regions, the supporting evidence that he presented was from cognitive semantic and psycholinguistic studies in general – no cross-linguistic, comparative data were taken into consideration. By incorporating insights from Gärdenfors's conceptual modeling technique, the CN approach aims to uncover and represent universals and variations in relation to how different languages encode a particular set of meanings by referring to cognitively informed principles, such as prototypicality and convexity. Indeed, there is a gap to be filled in the semantic typological literature regarding the application of geospatial techniques to explaining the organization of linguistic functions. In its origins, the technique of Voronoi tessellation finds its applications in many branches of studies, particularly those that involve cartographic methods, such as geography, climatology, ecology, and epidemiology. In linguistics, this technique is mainly applied in geolinguistic and dialectometric analyses of regional variations, e.g. Nerbonne (2009) and Goebl (2010). Such analyses generally involve geometric input, and the Voronoi structure is superimposed onto a geographical map to produce an interpretable visualization of the data. When applied to semantic typology, however, such a diagrammatic representation is not intended to convey any geospatial information, but rather information relating to the conceptual relations between a particular set of linguistic functions (meanings/uses).

Furthermore, apart from prototype points and their corresponding convex regions, a model in the CN approach also comprises a number of possible connections. A connection is basically a line linking any two points that are possibly related, and can be produced by a geometrical method called the Delaunay triangulation, which is also widely applied as a cartographic technique in geospatial analysis. Simply put, any three points make a Delaunay triangle if the circumcircle that passes through them is 'empty', i.e. does not contain any other function point within it. As illustrated in Figure 12, Points X, Y, and Z satisfy the Delaunay triangulation condition, i.e. no other point is contained within the circumcircle that passes through them. In other words, they are contained in three Voronoi regions spatially adjacent to one another. On the other hand, the circumcircle that passes through Points V, W, and Z does not have an empty interior, as it contains Points X and Y within it. So, Points V, W, and Z do not form a Delaunay triangle. In other words, the Voronoi regions that contain Points V, W, and Z are spatially adjacent to one another. As such, the Delaunay triangulation can be deemed to be the dual graph of the Voronoi tessellation, since the perimetric lines of the Delaunay triangles in a given domain correspond to the lines connecting each pair of the functions points with adjacent Voronoi regions. Considering that each of the triangles suggests a possible conceptual relation between the functions examined, a collection of such triangles of a particular set of function points in a given space forms a network of possible connections. Figure 13 shows possible connections between the functions in our constructed example, generated by means of Delaunay triangulation. In the figure, the conceptual proximities between each pair of function points are shown along the corresponding connecting lines.

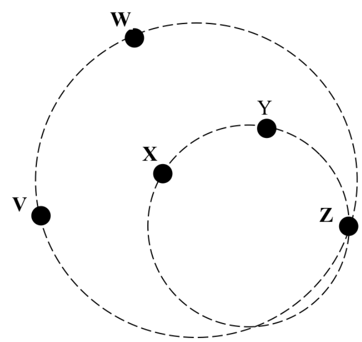

Figure 12 A Delaunay triangle (X-Y-Z) and a non-Delaunay triangle (V-W-Z), with circumcircles passing through the points

However, it should be noted that not every possible connection constitutes an optimal route. This is because, in some cases, the conceptual distance of a particular possible connection, i.e. a connection between a pair of function points with adjacent regions, is so great that it becomes improbable. Improbability here can be explained from various perspectives: psychology, historical, and typological. Psycholinguistically, an improbable connection arises when a direct association between a pair of word concepts is theoretically possible, i.e. not blocked by the presence of another intermediate concept, but requires too much mental effort to be cognitively realistic. A parallel situation can be discerned in historical linguistics when, although a particular pathway of semantic change can be posited in principle, there is no, or not enough, evidence in support of such a change. Moreover, in linguistic typology, an improbable connection, though predicted by the method, exhibits relatively little statistical evidence across languages. From any of the perspectives, such a connection tends to be avoided in the construction of a model, and some alternative, more probable ways to relate any two entities are posited. This process of identifying which connection is probable or improbable guarantees that the information conveyed by the spatial representation is statistically constrained. That is, statistical significance is conceptually translated as the probability of a connection.

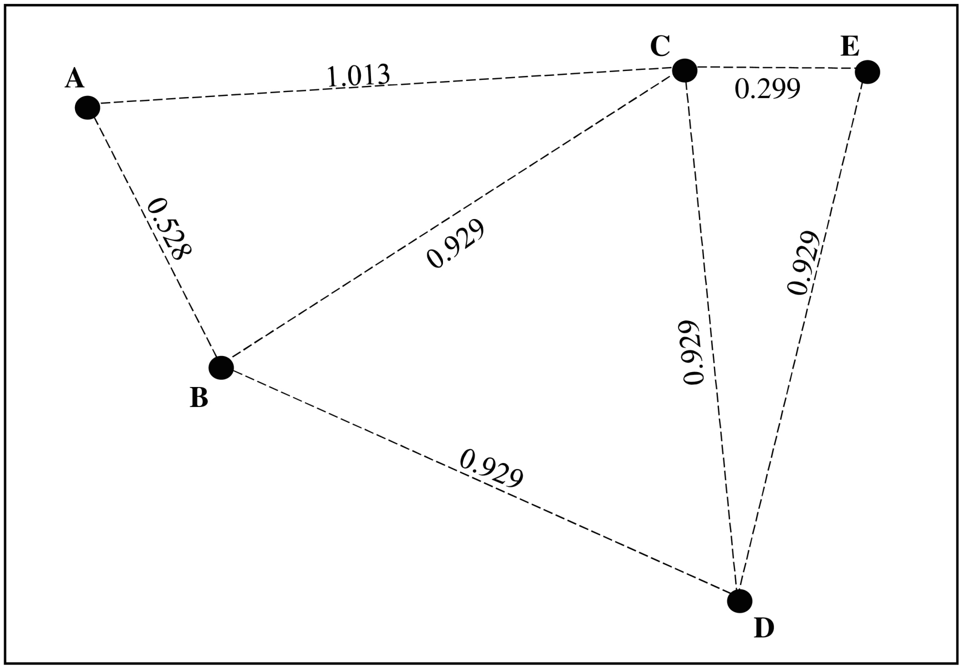

Figure 13 Function points A-E and possible connections between them, generated by means of Delaunay triangulation; all connections represented with proximity values

Furthermore, in the CN model, in which proximity between function points is conceptually interpretable, the probability of a connection can be determined based on its distance relative to those of the otherconnections belonging to the same Delaunay triangle. Specifically, as illustrated in Figure 14, for any set of connections that form the three sides of a scalene triangle, an improbable connection is one that makes the longest side. This is because the function points at the two ends of this connection are so conceptually far apart that the direct association between them is cognitively unrealistic. More probably, the two function points are linked indirectly via the third function point, i.e. along the two less distant connections. In the cases of equilateral and isosceles triangles, however, every connection is potentially probable. This is because each connection in such a triangular structure represents a geometrically optimal way to link the two function points. It should be noted, nevertheless, that in some cases, especially when a large number of data points are involved, the criterion for identifying a probable connection can be adjusted. That is, if the direct connection between a pair of function points that makes the longest side of a scalene triangle is not significantly longer than the connection that makes the second longest side, then it might be possible that the direct connection is considered to be an optimal route to link the two function points. In such a case, cognitive economy in terms of distance of connections is traded off for cognitive economy in terms of number of connections.

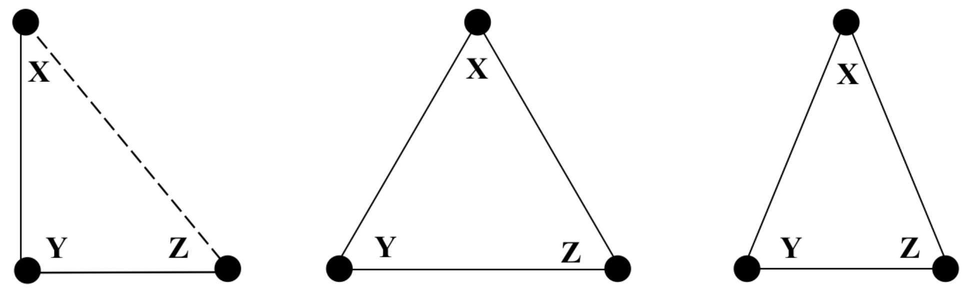

Figure 14 Possible connections between Points X, Y, and Z in different types of triangles: scalene (left), equilateral (middle), and isosceles (right) triangle. Solid lines represent probable connections and dotted lines represent improbable connections.

Let us now turn to the (im-)probability analysis of the connections in our constructed example, as shown in Figure 15. It should be noted that there are three Delaunay triangles in the network. A–B–C, B–C–D, and C–D–E. First, Triangle A–B–C is scalene. Although the direct connection between Points A and C is possible in principle, it is improbable statistically. This statistical improbability is represented geospatially as the longest side of the scalene triangle. Indeed, according to the data, there is no single construction that encodes Functions A and C without encoding Function B. However, as the other two connections of the triangle, i.e. A–B and B–C, are statistically valid, then it is probable that Points A and C are connected only indirectly via Function B. The second triangle, i.e. B–C–D, is an equilateral triangle, and so the connections that make its three sides, i.e. B–C, B–D, and C–D, are all probable in principle. Consequently, the connection that is shared by both and only the first and second triangles, i.e. B–C, can be confidently confirmed to be probable. The last triangle. i.e. C–D–E, is an isosceles triangle. The connection that forms the shortest side, i.e. C–E, is certainly probable. The other two equidistant sides, i.e. C–D and D–E, are both probable connections, as they represent equally probable ways to connect Point D to either Point C or E. Consequently, again, the connection that is shared by both and only the second and third triangles, i.e. C–D, can be confidently confirmed to be valid.

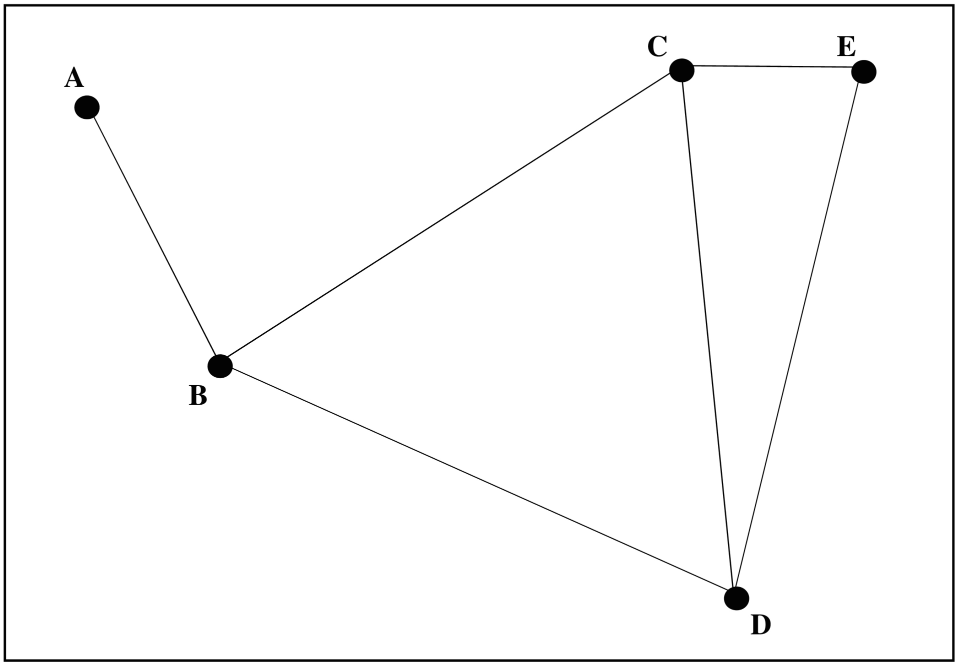

Figure 15 Network of probable connections between Points A-E.

A network of probable connections, in other words, a CN, like this serves as a conceptual representation of the relations between a set of functions in a given conceptual space. This conceptual representation is cognitively realistic because it takes into account the factors of mental association and effort. The CN of our constructed example, which consists of six probable connections (i.e. A–B, B–C, B–D, C–D, C–E, and D–E), makes a number of predictions. For instance, it predicts that it is impossible for a construction to encode Functions B and E without encoding Functions C and/or D. Also, it predicts that it is not probable for a construction to encode Functions A and C without encoding Function B. Furthermore, information on the functional distribution of a particular construction under investigation can be represented in the form of a subnetwork of probable connections – in other words, a conceptual route, drawn onto a given CN. To illustrate this, Figure 16 shows the conceptual routes of Constructions 1–3. The information on the probability of connections conveyed by a CN model is potentially more or less comparable to the implicational hierarchy posited by an SM model. To illustrate this, Figure 17 shows a model built using the traditional SM approach based on the data of our constructed example. It should be noted that the predictions formulated by the SM model exactly correspond to the predictions made based on the probable connections of the CN model. Therefore, it is apparent that the CN approach presents a mathematically formalized way of constructing a conceptual model with components that can be geometrically interpreted as absolute and/or statistical universals. As such, the CN model is preferable to the MDS model, which lacks principled ways to formulate implicational hierarchies, and to the SM model, the representation of which lacks computational tractability and statistical interpretability

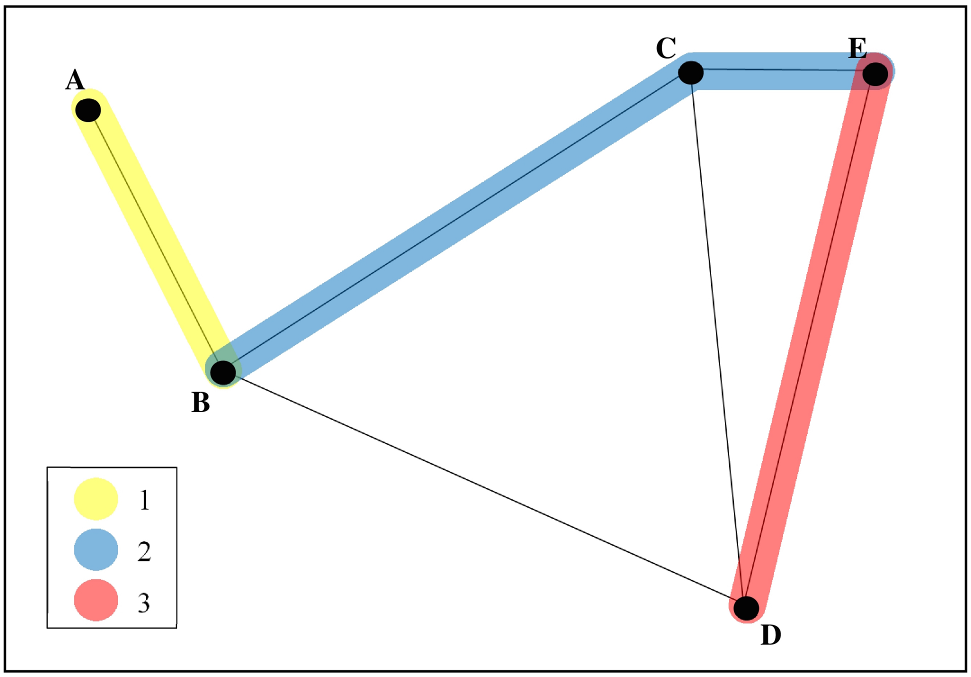

Figure 16 Conceptual routes of Constructions 1-4.

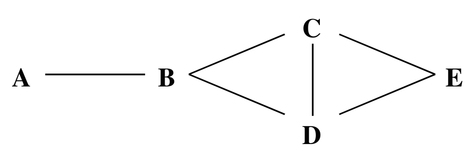

Figure 17 Traditional SM model based on the data of our constructed example.

6 The CN of indefinite pronouns

Now, let us apply the CN approach to the analysis of actual data, and then assess how well CN fares compared to the two previous approaches to semantic typology. This analysis is based on Haspelmath's (1997) data of the distribution of 139 indefinite pronoun constructions from 40 languages across nine functions. The 1,251 data points are in a binary code, with 1 meaning that a specific construction encodes such and such a function, and 0 meaning that it does not encode such and such a function, or simply that the data point is unavailable. Then, a Pearson correlation coefficient analysis, which is used to measure the linear dependence of the distributional patterns of each pair of functions, is performed on the table of distribution. The product of this analysis is a matrix of pairwise comparisons, with a larger value meaning a higher degree of correlation, and a smaller value meaning a lower degree of correlation, as shown in Table 5. Next, these values are transformed into conceptual proximities by converting similarities into dissimilarities, setting the smallest of the dissimilarities equal to 0, and then normalizing the values. Table 6 shows the matrix of transformed proximities between the indefinite pronoun functions. Furthermore, the MDS program PROXSCAL is applied on the matrix of normalized transformed proximities in order to determine the optimal coordinates of the points representing the functions investigated, as shown in Table 7, and to construct the representation of a conceptual space populated by those points, as shown in Figure 18. Finally, a stress test is performed on the model in order to evaluate its fitness and to determine the optimal dimensionality. Figure 19 shows the scree plot of the degree of fitness by the number of dimensions. It is apparent that the data can be optimally represented at two dimensions, with a stress level of .009 and a degree of improvement of .050.

At this point, let us tackle the problem of descriptive power by evaluating how well the proximity scaling method applied in the CN approach fares in comparison with the practice of manual counting in conventional semantic mapping and the Optimal Classification nonparametric unfolding algorithm applied in the MDS approach. First, both the CN and MDS approaches rely on a computationally tractable algorithm to automate a conceptual model. With the help of a statistical method or device, a high level of accuracy can be achieved even when a large sample size is involved. An SM model, on the other hand, is constructed manually. Therefore, not only is the process of constructing a model of this type time-consuming and prone to human error, but it is also difficult for a large amount of data to be accommodated. A second issue concerns goodness-of-fit. In the CN and MDS approaches, methods for assessing fitness can be applied in order to determine how well a

	SPKN	SPUN	IRNO	QUES	COND	INNE	COMP	DINE	FRCH
SPKN	.								
SPUN	.761	.							
IRNO	.446	.648	.						
QUES	.184	.335	.654	.					
COND	.184	.335	.654	.857	.				
INNE	-.158	-.151	.096	.443	.386	.			
COMP	-.452	-.434	-.233	.085	.173	.420	.		
DINE	-.345	-.354	-.232	-.047	-.163	.373	.054	.	
FRCH	-.340	-.360	-.275	-.226	-.101	.003	.566	-.141	.

Table 5 Matrix of pairwise comparisons of nine functions of indefinite pronouns, measured by means of the Pearson correlation coefficient

	SPKN	SPUN	IRNO	QUES	COND	INNE	COMP	DINE	FRCH
SPKN	.000								
SPUN	.124	.000							
IRNO	.494	.257	.000						
QUES	.802	.625	.250	.000					
COND	.802	.625	.250	.011	.000				
INNE	1.204	1.196	.906	.498	.565	.000			
COMP	1.550	1.529	1.293	.919	.815	.525	.000		
DINE	1.424	1.435	1.291	1.074	1.210	.580	.955	.000	
FRCH	1.418	1.442	1.342	1.284	1.137	1.015	.353	1.184	.000

Table 6 Matrix of transformed proximities between indefinite pronoun functions (matrix conditional, interval transformation)

	X	Y
SPKN	-.828	-.132
SPUN	-.734	-.055
IRNO	-.458	.001
QUES	-.166	.060
COND	-.152	-.027
INNE	.368	.227
COMP	.726	-.225
DINE	.549	.729
FRCH	.696	-.619

Table 7 Final coordinates of indefinite pronoun functions in a two-dimensional space

model lends itself to data representation. Also, the choice of dimensionality can be made on the basis of statistical principles. On the other hand, constructing an SM model involves, in essence, connecting any two functions that exhibit a perfect match. Any two functions are considered to match perfectly if none of the constructions examined encodes one but not the other. In other words, every pair of functions is initially hypothesized as connected, until there is negative evidence that refutes the connection. This kind of perfect match requirement imposed by the SM method is disadvantageous in that only possibilistic, as opposed to probabilistic, information can be represented in such a model. Moreover, the analyst's evaluation of goodness-of-fit and their decision regarding the number of dimensions needed to represent the data are probably not statistically informed.

Furthermore, there is a point of departure that distinguishes CN from Croft and Poole's (2008) MDS approach. That is, the CN approach makes use of a mathematical method that offers a more precise way to determine the conceptual proximities between the linguistic functions. To illustrate the case in point, let us take a look at how a particular problematic function point (i.e. indirect negation) is positioned differently in the two approaches. In Croft and Poole's (2008) MDS model of indefinite pronouns, displayed in Figure 5, the indirect negation function point is positioned undesirably closer to the conditional function point than to the question function point, despite the empirical evidence that there is no single construction that encodes the conditional function without encoding the question function. This undesirable

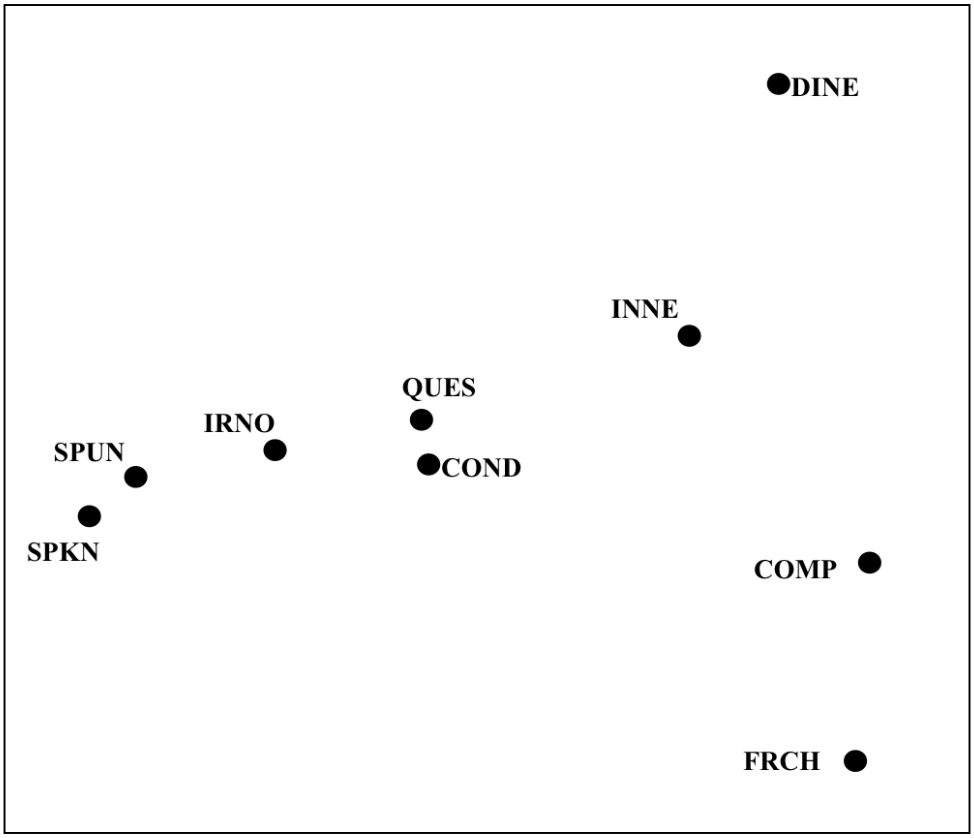

Figure 18 Points of indefinite pronoun functions plotted onto a two-dimensional space.

positioning admittedly results from the insufficiency of data. However, with the same amount of data, a higher level of accuracy is achieved in the CN model. As shown in Figure 17, the indirect negation function point is now positioned closer to the question function point than to the conditional function point. This alternative positioning is performed by PROXSCAL so as to reflect the conceptual proximities between these functions as closely as possible. By applying such a statistical technique, I suggest that the spatial configuration of the CN model reflects the conceptual relations between the relevant functions in a more accurate fashion.

Let us now resume the process of constructing the CN model of indefinite pronouns. The final coordinates in Table 7 are used to plot function points in a geospatial representation, by using QGIS. Each point is considered to be an ideal site for the prototype of each function, contained in a convex

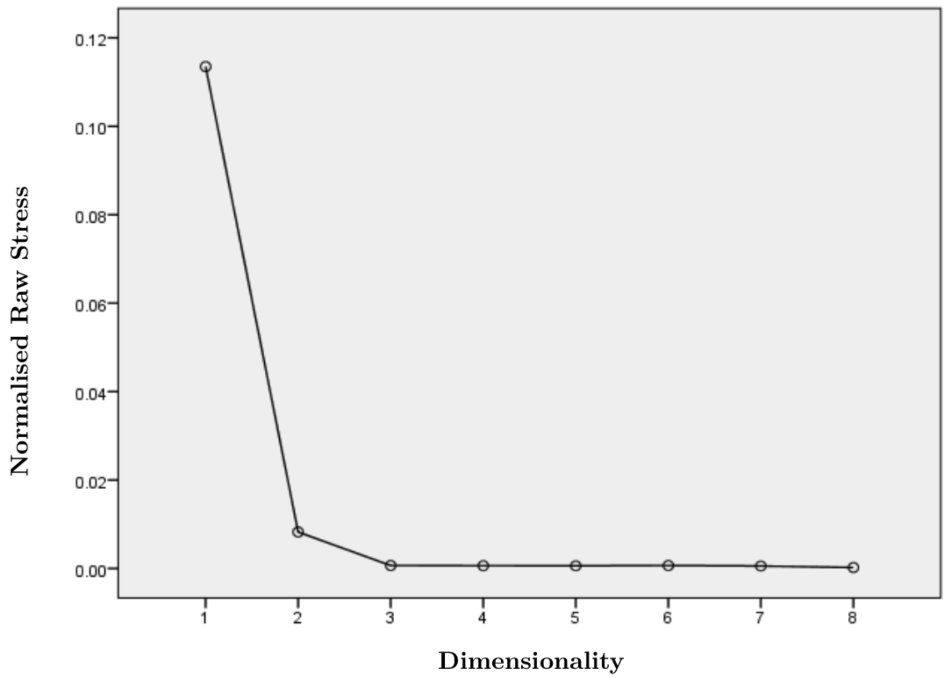

Figure 19 Levels of normalized raw stress by the number of dimensions.

region representing the conceptual subspace of that particular function. The partitioning of the conceptual space into convex regions is performed using the Voronoi tessellation technique available in PROXSCAL. Conceptually, each region can be deemed to represent the set of intersecting properties or parameters that make up that particular function, and functions with adjacent regions are considered to be immediately related. Figure 20 shows the prototype points of the nine indefinite pronoun functions, each contained in its corresponding convex region. Next, connections are established between any pair of prototype points whose regions are adjacent, using the technique of Delaunay triangulation. Potentially, these connections represent possible conceptual relations between the functions, as shown in Figure 21. However, not every connection that is possible in theory is statistically probable. In order to determine which possible connection is probable, the optimal route analysis is performed. Probable connections are basically those that do not constitute the longest side of a scalene triangle. Potentially, a CN that takes into account probabilistic information like this can represent how linguistic functions are organized and processed in a cognitively realistic fashion. Figure

22 shows such a CN, representing the indefinite pronoun functions and the probable connections between them. Finally, at the intra-linguistic level, the functional range of a particular construction can be represented as a route connecting prototype points within the network. To illustrate, the conceptual routes of four indefinite pronoun constructions with *-va, vre- -un, ni-*, and *ori*-in Romanian are diagrammatically represented in the CN in Figure 23.

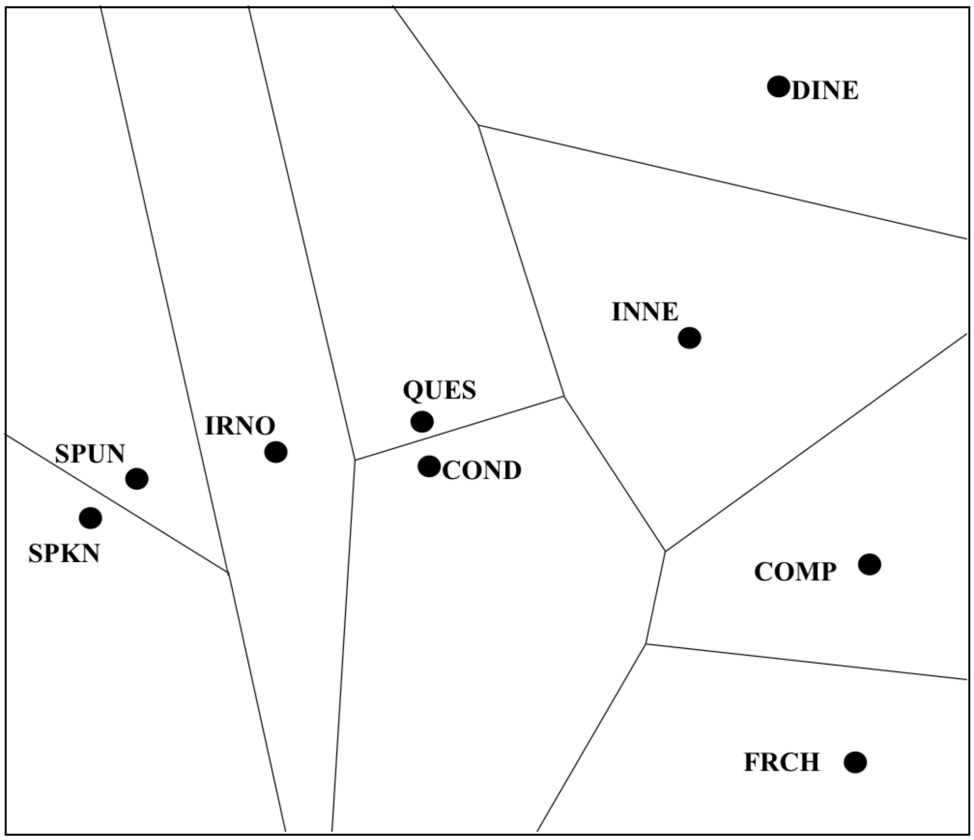

Figure 20 Prototype points of the indefinite pronoun functions and their corresponding convex regions, generated by means of Voronoi tessellation.

At this point, let us now make an assessment of the CN approach in comparison with the SM and MDS models in terms of explanatory power. It should be noted that the latter two models constitute two opposite extremes, i.e. probability-free and constraint-free models, while the CN falls somewhere in between. In a probability-free model like SM, the configuration of components is so constrained that it produces a discrete output. That is, the objective

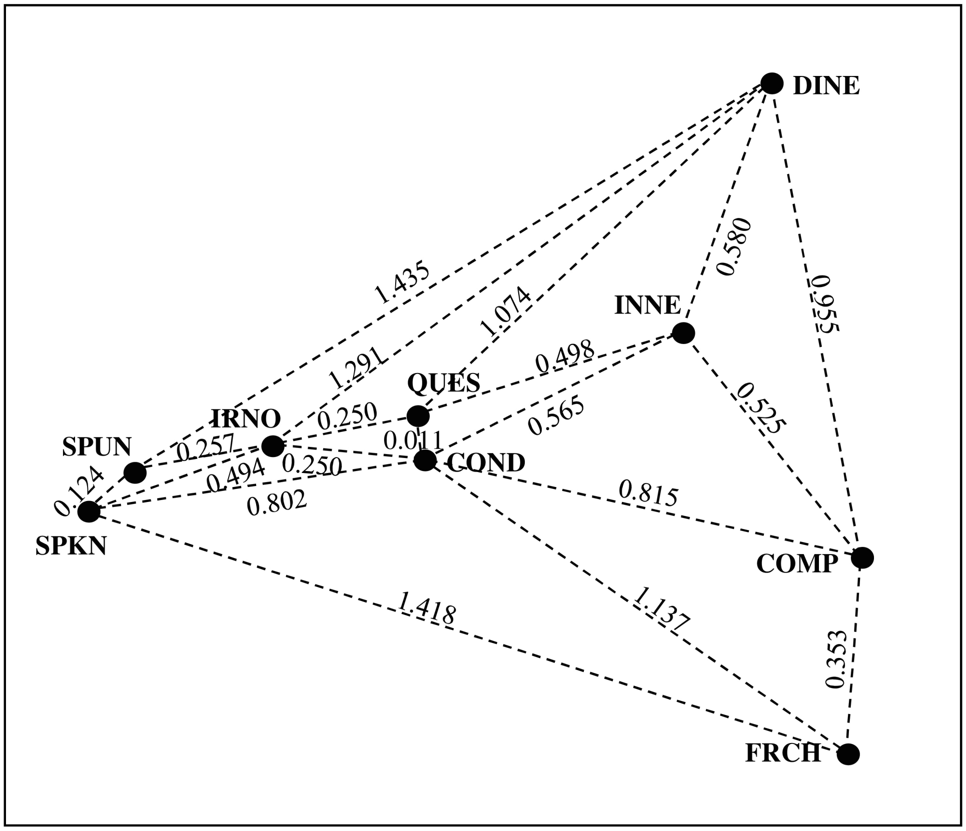

Figure 21 Indefinite pronoun function points and possible connections between them, generated by means of Delaunay triangulation; all connections represented with proximity values.

of an SM model is to formulate absolute, implicational universals by finding perfect matches between functions. Any two functions can be deemed to match perfectly if, when these two functions are encoded by a particular construction, there is no other function that is encoded by that construction. Two such perfect matching functions are represented as connected to each other. This criterion of perfect matching is tested and retested in a particular set of functions against a class of cross-linguistically comparable constructions. The result, then, is an implicational map that predicts the encoding behavior of every possible construction that belongs to that class. Two possibilities are represented in such a map, i.e. either the connection is possible or it is impossible. The spatial configuration of an SM model, however, does not convey any probabilistic information. For example, in Haspelmath's (1997)

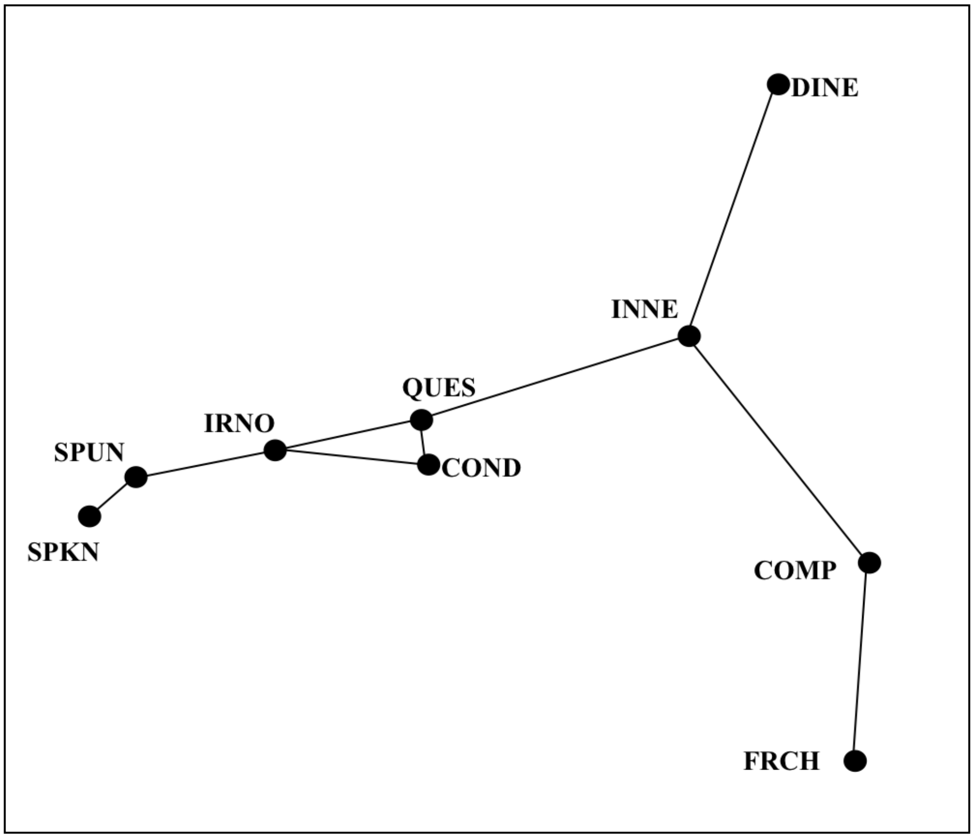

Figure 22 CN of probable connections between indefinite pronoun function points.

model of indefinite pronouns, as shown in Figure 2, the comparative function is represented as connected to three functions, i.e. the conditional, indirect negation, and free choice functions, with lines of an equal length and at positions of an equal distance from the comparative function.

However, a simple statistical measure like the Pearson correlation coefficient can be used to show that in fact there are significant differences between the conditional, indirect negation, and free choice functions in relation to the comparative function, in terms of patterns of encoding behaviors. Simply put, the degree of correlation can be used to measure how much the encoding behaviors of two functions exhibits linear dependence. Specifically, if two functions tend to behave similarly, the correlation is positive, but if they tend to behave differently, the correlation is negative. It is found that the pattern of encoding behavior of the comparative function is most similar to that of the

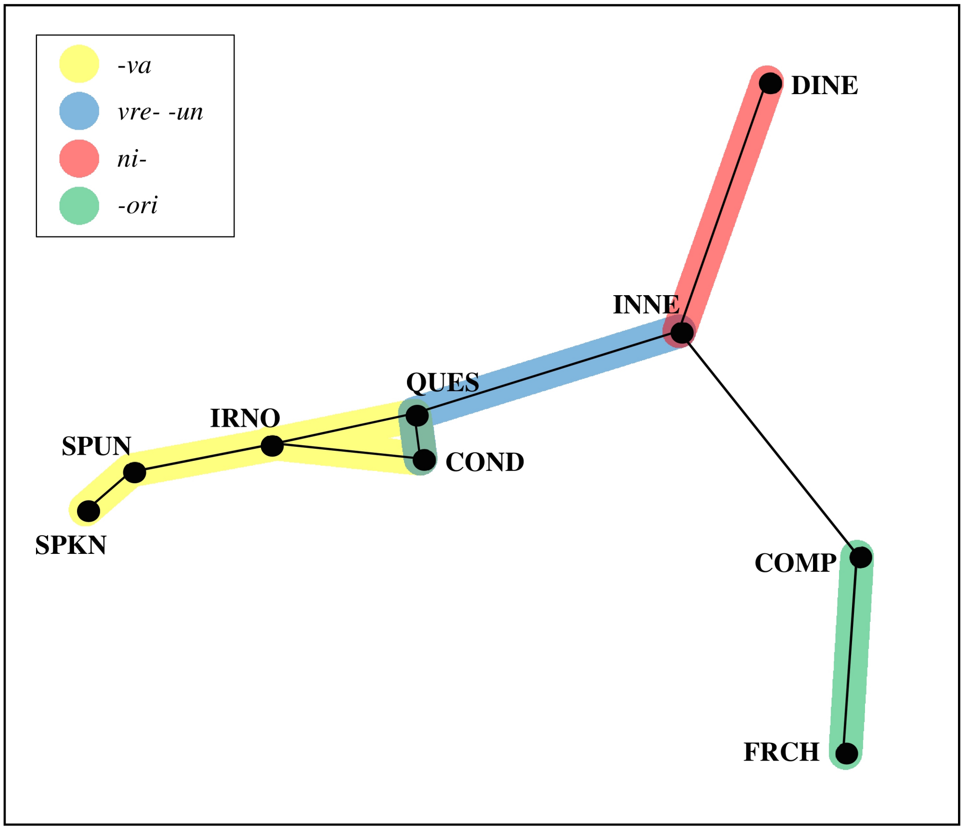

Figure 23 Conceptual routes of four Romanian indefinite pronouns.

free choice function, with a correlation value of .566. Somewhat less similar, however, is the encoding behavior pattern of the comparative function to that of the indirect negation function, with a correlation value of .420. Furthermore, the comparative function behaves least like the conditional function, with a correlation value of .173. In simple terms, these different correlation values indicate different degrees of probability that those functions will be encoded by the same construction. This kind of information about different degrees of probability, however, is not translated by the practice of semantic mapping at all, either in absolute or relative forms. Worse still, it is doubtful if such information can be accommodated at all by an SM model, given the current method.

On the other hand, a model that takes into account conceptual proximities between the functions, like an MDS and CN model, does not suffer from this problem. Specifically, in an MDS or CN model, geometric distances between the functions examined are intended to reflect the probability that they will be encoded by the same construction, such that a pair of constructions with a higher probability in this regard are positioned closer to each other, and a pair of constructions with a lower probability are positioned less close to each other. As such, the CN method can potentially be used to construct a model that better represents the encoding behaviors of related functions. For example, in the CN model of indefinite pronouns, the proximities between these functions are different, i.e. .353 between the comparative and free choice constructions, .525 between the comparative and indirect negation functions, and .815 between the comparative and conditional functions. From this example, it is clear that a non-probability-free model, like CN, is more advantageous than a traditional SM model in that it can represent the encoding behaviors of constructions examined in a more informative fashion.

Moreover, the incorporation of probabilistic information in the CN approach makes it a more efficient tool for semantic typology. Specifically, the approach distinguishes between possible and probable connections between functions. For example, it is found that although the direct connection between the conditional and comparative functions is possible, it is not a probable one. In fact, of the 32 constructions that encode these two functions, only two do so without encoding the question and indirect negation functions. For instance, as shown in Figure 24, the functional range Finnish *hyvänsä*-series encompasses three functions (conditional, comparative, and free choice), thus violating the probabilistic connectivity principle. The other 30 constructions, on the other hand, encode the question and/or indirect negation functions in addition to the conditional and comparative functions. In other words, it is significantly more probable that the latter two functions are conceptually connected via the former two functions. However, the traditional SM model relies on the discrete notion of perfect fit, and thus allows every instance of connection to be counted towards constituting a linguistic universal. As such, the model potentially suffers from the problem of overprediction. That is, 129 different patterns of functional distribution can be derived from Haspelmath's (1997) SM model of indefinite pronouns, while in fact only 39 patterns are attested to in the data. So, about 70% of the patterns are overpredicted. The CN model, on the other hand, makes a more realistic representation of the universals and variations, as it predicts 83 different patterns of functional distribution, thus reducing the rate of overprediction to 53%.

However, a constraint-free model, like an MDS model, is also problematic. Specifically, in an MDS model, functions are represented as points in a multidimensional space such that the distances between these points are intended to represent the conceptual distances between the functions. Although this kind of

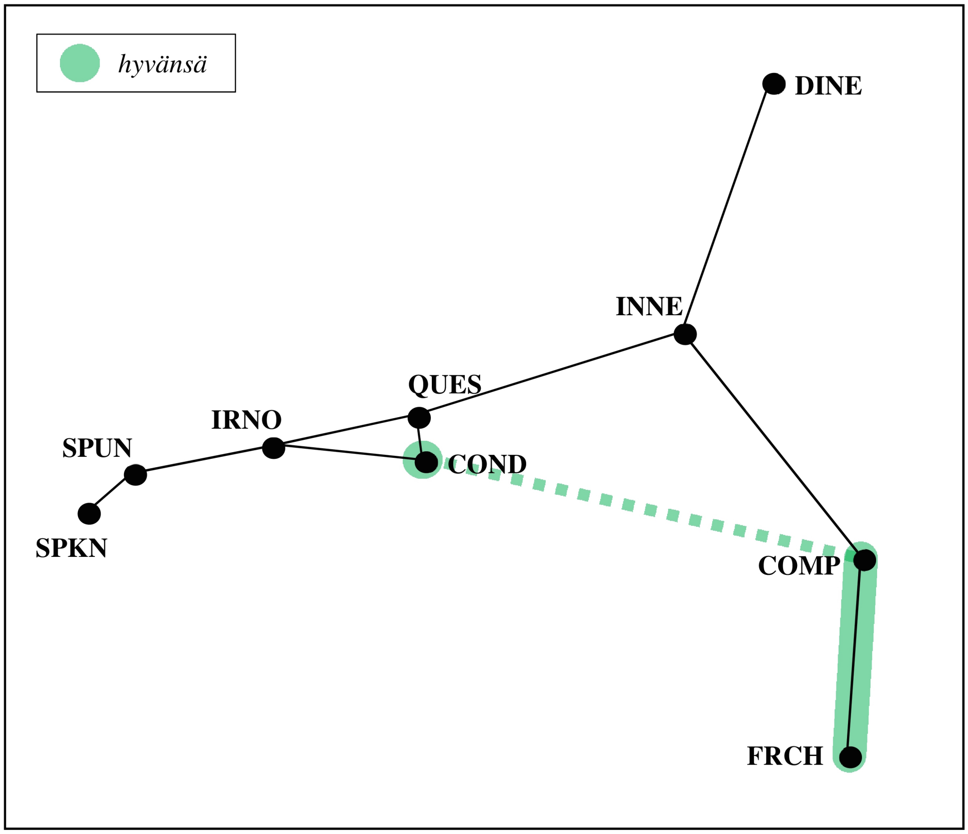

Figure 24 Conceptual route of Finnish *hyvänsä*-series.

representation is useful in the sense that it allows probabilistic information to be maximally conveyed, it does not possess much explanatory power compared to SM and CN. A traditional SM model explicitly represents an implicational map that can be used to explain how any construction that belongs to the class should behave in terms of function encoding. To illustrate this, in Haspelmath's (1997) model of indefinite pronouns, it is explicitly predicted that every construction that encodes direct negation and comparative functions must also encode the indirect negation function. Such prediction is important in that it makes the model testable, and thus falsifiable, by empirical evidence. For example, if it is later found that there exists a construction in such and such a language that can encode the direct negation and comparative functions without encoding the indirect negation function, then Haspelmath's model is falsified, and will probably be modified in order to better accommodate the data.

Similarly, in a CN model, constraints also exist, but in the form of the network of possible and probable connections, which is intended to capture both construction-specific add language-specific variations, and typological universals. In other words, a network constructed using this approach conveys information at both the inter-linguistic and intra-linguistic levels. Inter-linguistically, it represents a potential model of how a set of linguistic functions (meanings/uses) can be related conceptually. Intra-linguistically, this network makes predictions as to how any construction should behave in terms of function encoding. For example, the CN model of indefinite pronouns predicts that a direct link between the direct negation and comparative functions without including an indirect negation function is not probable, i.e. the direct negation and comparative functions do not form a cluster, as this link violates the probabilistic connectivity principle. However, if empirical evidence states otherwise, e.g. that, in fact, cross-linguistically there exist a substantial number of constructions that encode the direct negation and comparative functions without encoding the indirect negation function, then the model is falsified and needs modification. Actually, this kind of prediction is quite similar to that made by the SM model, except that the former additionally conveys probabilistic information while the latter does not.

An MDS model, however, does not seem to have such explanatory power. Specifically, there exists no prediction concerning the encoding behaviors of constructions in any explicit form. Such a model consists of points representing functions in a conceptual space, with no connections made between any functions. Given that the encoding behavior of a particular construction can be represented by a cutting line drawn onto the conceptual space, it might be possible that some predictions can be made by the configuration of cutting lines. Nevertheless, there simply exist too many possible patterns of function encoding, some of which might not seem very likely. For example, a cutting line that separates the space into two parts, including the direct negation and comparative functions, without including the indirect negation functions can be drawn conveniently onto the space, representing a construction that encodes only the former two constructions. Empirical evidence so far indicates, however, that there exists no such construction. In fact, every construction that encodes both the direct negation and comparative functions also encodes the indirect negation. The absence of such a construction cannot be inferred from the model.

A final remark on the assessment of the three approaches to semantic typology concerns discreteness and typological features. By design, the MDS model is non-discrete, as it has no principled way of partitioning the conceptual space into different subspaces, representing different values a particular binary feature. The traditional SM model, on the other hand, is essentially discrete and can accommodate the representation of typological features. That is, it is required that 'the implicational map should be arranged in such a way that all functions that share some relevant characteristics [...] form a contiguous area on the map' (1997). Likewise, the CN model, which derives discrete information from non-discrete information, also has a way of accommodating the representation of binary features. Specifically, if a set of functions share the same value in respect of a particular typological feature, the convex regions that represent them must be adjacent, thus forming a contiguous subspace. For example, according to Haspelmath, four binary features can be used to characterize the nine functions of indefinite pronouns: (i) known to the speaker vs. unknown to the speaker; (ii) specific vs. non-specific; (iii) scalar endpoint v. no scalar endpoint; and (iv) in scope of negation vs. not in scope of negation. In addition, a sub-feature of (iv), reversed vs. non-reversed scale, can be applied to the set of functions with a scalar endpoint. Figures 25–28 represent the subspaces of these typological features, which are successfully accommodated in the CN model.

7 Conclusion

At this point, let us conclude the paper by summarizing the main points made and suggesting directions for further research. First, defined as a systematic cross-linguistic study of how meaning is encoded by linguistic signs, semantic typology can be achieved by various approaches. One such approach is the SM approach, which aims to build language-specific constructional matrices and map them onto a universal conceptual space. Represented in the conceptual space are nodes representing distinct functions. These function nodes are positioned in such a way that any two functions encoded by the same linguistic form without implicating the encoding of any other function are represented as contiguous. A notable application of the SM approach is Haspelmath's (1997) model of indefinite pronouns. The SM model, however, has inherent drawbacks. As the model relies on the principle of perfect fit in postulating language universals in the form of implicational hierarchy, it does not take into account probabilistic information. Furthermore, as the model cannot be automated using a computationally tractable algorithm, it does not lend itself to a large-scale analysis with a great number of data points involved. To overcome these advantages, Croft and Poole (2008) proposed a new approach to semantic typology called the MDS approach. A model built using this approach basically consists of a number of cutting lines representing the functional ranges of different constructions investigated. Each of these lines represents a classification of the functions that belong to a given domain by cutting

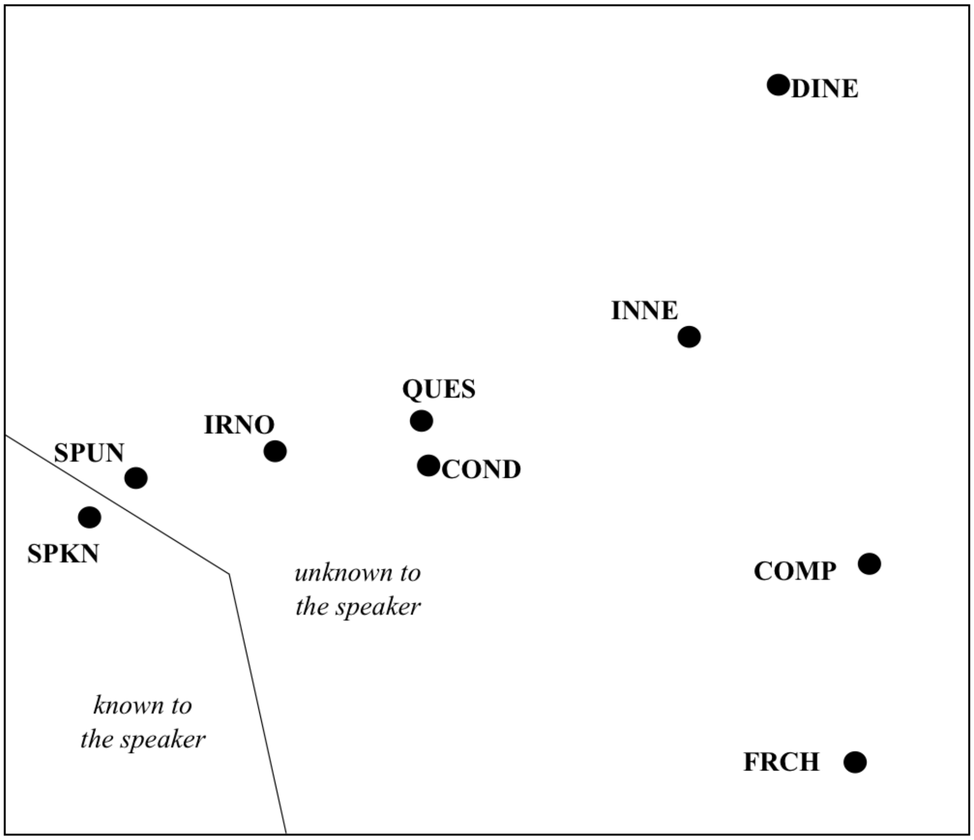

Figure 25 Subspaces of feature known vs. unknown to the speaker.

the conceptual space into two regions: that is, the region including only the functions encoded by a particular construction (the Y region) and the region including only the functions not encoded by it (the N region). Moreover, the intersection of these cutting lines forms a polytope in which the ideal point of each function is located. Then, a measurement of fitness is performed on the model to determine the level of the correct classification of the data.

However, the MDS model still suffers from a number of disadvantages. Most notably, it lacks both descriptive and explanatory power. Specifically, the distances between the ideal points do not seem to represent the conceptual proximities between the different functions in an accurate fashion. This is because for an ideal point to be accurately positioned, it requires that the corresponding polytope be sufficiently small. For most studies on semantic typology, however, the high level of accuracy cannot be achieved, as relatively small data sets are involved, hence somewhat large polytopes. Furthermore,

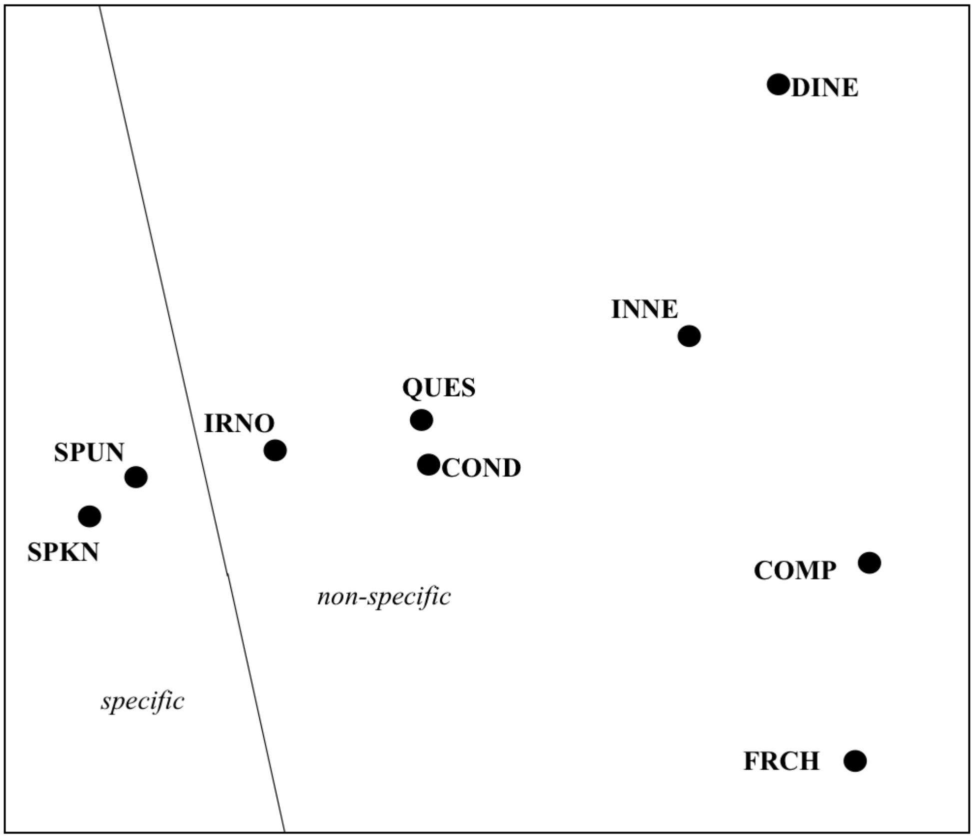

Figure 26 Subspaces of feature specific vs. non-specific.

the MDS model does not offer a principled way to determine different levels of relations between functions. As a result, the model cannot accommodate the representation of language universals in the form of implicational hierarchy. In response to these inadequacies of the previous models, I proposed a new approach called the CN approach. In this approach, a high level of descriptive accuracy can be achieved as the conceptual proximities between the functions examined are determined by a matrix of pairwise comparisons of the Pearson correlation coefficient values. These values are then used to determine the optimal coordinates of the prototype points. Moreover, the fitness of a CN model can be measured by assessing how accurately the conceptual proximities between the functions are translated into the geometric distances between the points. Then, by using the geospatial technique of Voronoi tessellation, the conceptual space is partitioned into a number of convex regions, each with its corresponding prototype point positioned at the center. Each pair

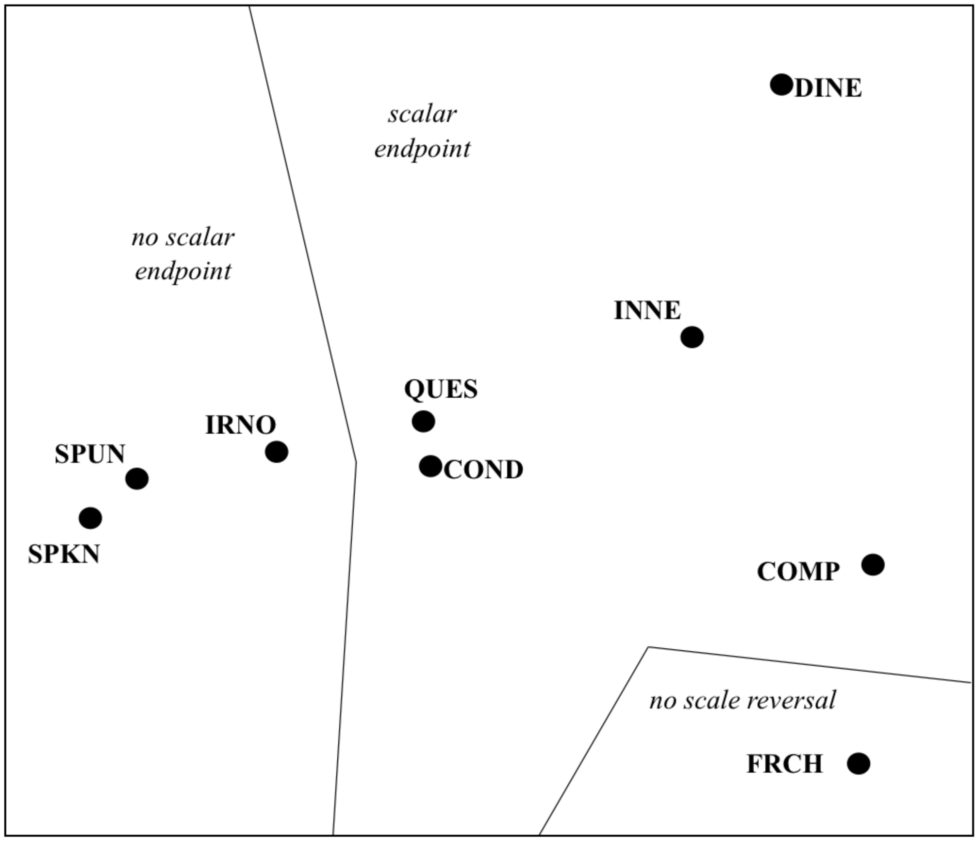

Figure 27 Subspaces of feature scalar endpoint vs. no scalar endpoint and pf sub-feature scale reversal vs. no scale reversal.

of points and regions potentially represents the conceptual structure of its corresponding linguistic function. Furthermore, the Delaunay triangulation technique is applied to create a network of possible connections between the functions, such that any two points with adjacent regions are represented as having a connecting line between them. However, not every possible connection is statistically probable, and so a probabilistic analysis is performed on the different routes of the network. Basically, a connection between any two points constitutes a probable route if it does not form the longest side of a scalene triangle. This criterion is applied to guarantee that the functions are connected in a cognitively realistic fashion, i.e. by minimizing the mental effort and maximizing the ease of learning/processing. The final result is a network of probable connections, which can be interpreted as representing the language universals in the form of implicational hierarchy.

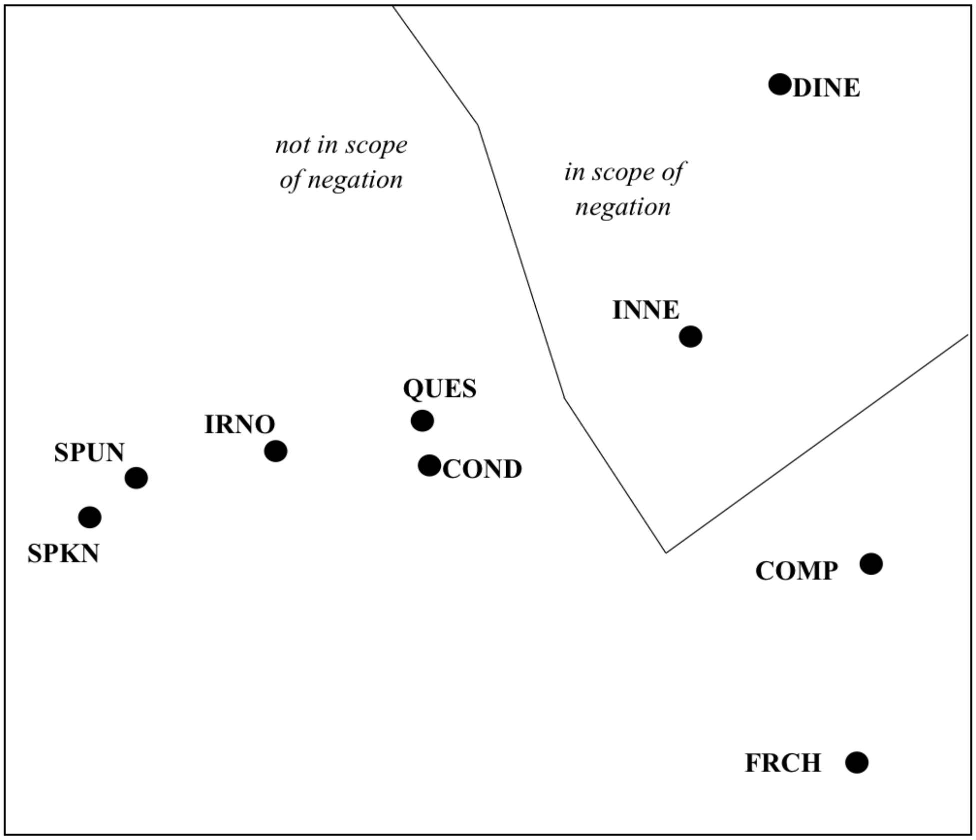

Figure 28 Subspaces of feature in vs. not in scope of negation.

Also, this paper has presented an application of the CN approach to an action data set. This data set consisted of 1,251 data points, representing the distribution of 139 indefinite pronoun constructions across nine functions. It was found that the CN model of indefinite pronouns is better than the previously proposed SM and MDS models in that it is neither probability-free nor constraint-free. Specifically, the CN model can accommodate both non-discrete information, in the form of varying degrees of distance between function points, and discrete data, in the form of the distinction between probable and non-probable connections. All in all, as illustrated above, the CN model is descriptively more accurate, and achieves a higher level of explanatory power. However, to test whether the CN approach is robust against different types of data sets, and whether it can produce reliable and consistently good results, wider applications of the approach are needed. For example, it is suggested that the approach be applied to much larger data sets, e.g. van der Auwera and Plungian's (1998) modality data. As it is claimed that the approach is cognitively realistic, it would be interesting for a model constructed using this approach be tested for its validity against psychological experiments or neurological evidence. Last, it is likely that the approach is not limited only to semantic typological applications. It might potentially be applied to explain non-semantic data, non-typological data, or even non-linguistic data. For instance, it is possible that the CN approach can shed some light when applied to Poole's (2000) US parliamentary voting data.

References

Anderson, L. (1974). Distinct sources of fuzzy data: ways of integrating relatively discrete and gradient aspects of language, and explaining grammar on the basis of semantic fields. In R. Shuy, & C.-J. Bailey, *Towards tomorrow's linguistics* (pp. 50-64). Washington, D.C.: Georgetown University Press.

Berlin, B., & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution.* Berkeley: University of California Press.

Croft, W. (2001). *Radical Construction Grammar. Syntactic Theory in Typological Perspective.* Oxford: Oxford University Press.

Croft, W., & Poole, K. (2008). Inferring universals from grammatical variation: Multidimensional scaling for typological analysis. *Theoretical Linguistics*, 1-37.

de Haan, F. (2010). Building a Semantic Map: Top-Down versus Bottom-Up Approaches. *Linguistic Discovery 8:1*, 102-117.

Evans, N. (2010). Semantic Typology. In J. J. Song, *Oxford Handbook of Linguistic Typology* (pp. 504-533). Oxford/New YorK: Oxford University Press.

Evans, V. (2004). *The Structure of Time. Language, meaning and temporal cognition.* Amsterdam/Philadelphia: John Benjamins.

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought.* Cambridge, MA: MIT Press.

Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces.* Cambridge, MA: MIT Press.

Goebl, H. (2010). Dialectometry and quantitative mapping. In A. Lameli, R. Kehrein, & S. Rabanus, *Language and Space, Vol. 2: Language and Mapping* (pp. 433-457). Berlin: De Gruyter Mouton.

Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to*

Argument Structure. University of Chicago Press: Chigago.

Goldberg, A. (2006). *Constructions at Work: the nature of generalization in language.* Oxford University Press: Oxford.

Haspelmath, M. (1997). *Indefinite pronouns.* Oxford: Clarendon Press.

Kemmer, S. (1993). *The middle voice.* Amsterdam: Benjamins.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind.* Chicago: University of Chicago Press.

Langacker, R. (1987). *Foundations of Cognitive Grammar, Volume I: Theoretical Prerequisites.* Stanford, California: Stanford University Press.

Langacker, R. (1991). *Concept, Image, and Symbol: The Cognitive Basis of Grammar.* Berlin & New York: Mouton de Gruyter.

Langacker, R. (2008). *Cognitive Grammar: A Basic Introduction.* New York: Oxford University Press.

Lounsbury, F. (1964). The structural analysis of kinship semantics. In H. Hunt, *Proceddings of of the Ninth International Congress of Linguistics* (pp. 1073-1093). Mouton.

Nerbonne, J. (2009). Data-driven Dialectology. *Language and Linguistics Compass 3:1*, 175-198.

Poole, K. (2000). Non-parametric unfolding of binary choice data. *Political Analysis 8*, 211-237.

Saussure, F. (1916). *Cours de linguistique générale.* Lausanne/Paris: Payot.

Tyler, A., & Evans, V. (2003). *The Semantics of English Prepositions: Spatial Sciences, Embodied Meaning, and Cognition.* Cambridge: Cambridge University Press.

van der Auwera, J., & Plungian, V. (1998). Modality's semantic map. *Linguistic Typology 2.1*, 79-124.

Zwarts, J. (2010). Semantic map geometry: Two approaches. *Linguistic Discovery 8:1*, 377-395.

James N. Chancharu
Department of Theoretical &
Applied Linguistics, MML
Sidgwick Avenue
Cambridge, CB3 9DA

nc368@cam.ac.uk