

## ARTICLE

# The semantic map model: State of the art and future avenues for linguistic research

Thanasis Georgakopoulos<sup>1</sup>  | Stéphane Polis<sup>2</sup> 

<sup>1</sup>Service d'Égyptologie, Département des sciences de l'Antiquité, Université de Liège

<sup>2</sup>Service d'Égyptologie, Département des sciences de l'Antiquité, Université de Liège/F.R.S.-FNRS

**Correspondence**

Thanasis Georgakopoulos, Service d'Égyptologie, Département des sciences de l'Antiquité, Université de Liège  
Email: athanasios.georgakopoulos@uliege.be

**Funding information**

Marie Curie BeIPD postdoctoral fellowship from the European Commission; National Fund for Scientific Research (Belgium)

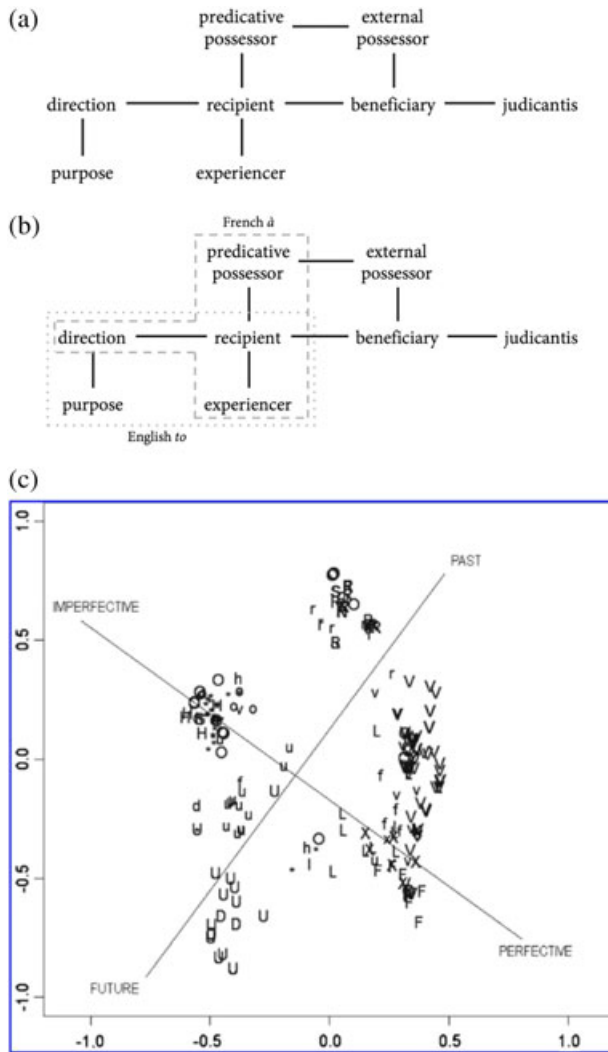
**Abstract**

The semantic map model is relatively new in linguistic research, but it has been intensively used during the past three decades for studying both cross-linguistic and language-specific questions. The goal of the present contribution is to give a comprehensive overview of the model. After introducing the different types of semantic maps, we present the steps involved for building the maps and discuss in more detail the different types of maps and their respective advantages and disadvantages, focusing on the kinds of linguistic generalizations captured. Finally, we provide a thorough survey of the literature on the topic, and we sketch future avenues for research in the field.

## 1 | INTRODUCTION: WHAT IS A SEMANTIC MAP?

This paper provides an overview of the semantic map model, a relatively new approach in linguistic research. The model has been intensively used during the past three decades for studying both cross-linguistic and language-specific questions. A semantic map is a way to visually represent the interrelationships between meanings<sup>1</sup> expressed in languages. One can distinguish two types of semantic maps: classical maps and proximity maps (van der Auwera, 2013; see Sections 4 and 5 respectively for alternative labels). Classical semantic maps typically take the form of a graph—with nodes standing for meanings and edges between nodes standing for relationships between meanings. Figure 1a is a textbook example of a classical semantic map for dative functions.

<sup>1</sup>Throughout this paper, we use the neutral term “meaning,” rather than the technical “signified,” or the less appropriate label “concept” sometimes found in the literature in order to refer to the nodes of the map. This term can refer to both coded and contextually inferred meanings (Ariel, 2008), and as such covers also “functions” and “uses.”



**FIGURE 1** (a) Semantic map of dative functions (adapted from Haspelmath, 2003, p. 213). (b) Semantic map of dative functions, with the areas covered by English *to* and French *à* (adapted from Haspelmath, 2003, pp. 213, 215). (c) Spatial model of tense and aspect with Dahl's prototypes (Croft & Poole, 2008, p. 26)

In such maps, two meanings are connected if they are expressed by the same linguistic item in at least one language. These maps are inferred from typological data, based on the hypothesis that language-specific patterns of polysemy<sup>2</sup> point to recurrent relationships between meanings across languages. Figure 1a shows, for instance, that the meanings “purpose” and “direction” are closely associated and predicts that, if a linguistic item expresses these two meanings and an additional one, it should necessarily be “recipient,” because it is the only meaning directly

<sup>2</sup>Polysemy refers to the phenomenon, whereby two or more related meanings are associated with a single lexical, grammatical, or even constructional item. In the literature, the terms “multifunctionality” or “polyfunctionality” are also used to refer to polysemous grammatical items (see, e.g., the use of the term “multifunctionality” by Haspelmath, 2003, in the context of semantic maps). For lexical items, François (2008) coined the term “colexification” to refer to “the capacity, for two senses, to be lexified by the same lexeme in synchrony” (François, 2008, p. 171).

connected to “purpose-direction.” The cross-linguistic regularities in semantic structure represented by semantic maps can be tested empirically and falsified by additional evidence (Cysouw, Haspelmath, & Malchukov, 2010, p. 1).

In order to visualize the meanings of language-specific items, one simply has to map them onto the graph. Figure 1b illustrates how this mapping works by including the boundaries of the English preposition *to* and the French preposition *à*: the two items share the meanings “direction,” “recipient,” and “experiencer,” but “purpose” is only expressed by *to* and “predicative possessor” by *à*. Furthermore, one can notice that they cover connected regions of the graph (see Section 2.2).

Proximity maps, on the other hand, are not graphs: the meanings or uses, represented by points, are distributed on a two-dimensional space using multivariate statistical techniques (usually Multi-Dimensional Scaling, MDS in short). The distance between the points of the map is indicative of their (dis)similarity, hence the label “proximity map.” Like classical semantic maps, proximity maps can be construed based on a semantic analysis of cross-linguistic data, but they may also be plotted on the basis of data alone (Narrog & van der Auwera, 2011, pp. 320–321). As such, they are a way to do “typology without types” (see Section 5).

Figure 1c, for instance, is based on a large dataset of tense-aspect constructions (Dahl, 1985). The points on the map are contexts of occurrence of prototypical tense-aspect clusters (H = Habitual; S = Habitual Past; O = Progressive; U = Future; V = Perfective, etc.), and the distance between any pair of dots reflects the probability that two contexts will be expressed by the same form in the languages of the sample. As can be observed, the points cluster rather well from a semantic point of view and can subsequently be analyzed along two axes: imperfective–perfective and future–past.

Details about the two different types of maps, their premises, and the generalizations that emerge from each of them will be given in different parts of this contribution. The paper is structured as follows. Section 2 discusses the basic principles underpinning the construction of classical semantic maps. Section 3 examines the usefulness of this approach in typology, semantics, and historical linguistics. Section 4 presents the different kind of representation techniques used in the literature for classical semantic maps and the types of knowledge that these representations capture. Based on a critical evaluation of the classical model, Section 5 introduces the proximity maps, which rely on an alternative plotting method and visualization technique. An overview of the literature on semantic maps is provided in Section 6, and we describe future avenues for the field in Section 7, focusing on the tools that allow an automatic plotting of classical semantic maps based on cross-linguistic polysemy data.

## 2 | HOW IS A CLASSICAL SEMANTIC MAP BUILT?

In order to describe the steps involved for building a classical semantic map, we take as point of departure Hjelmslev’s (1961, p. 54) famous example regarding the linguistic expressions of meanings belonging to the semantic field TREE–WOOD–FOREST, as articulated in Haspelmath (2003, p. 237). Looking at four languages, namely, Danish, French, German, and Spanish, the lexemes in this semantic field compare as in Table 1.

Table 1 shows that each language lays down its own boundaries at the semantic level (the *content-form* in Hjelmslev’s terminology). To put it otherwise, one observes a language-specific partitioning of the semantic domain by language-specific forms. The challenge for the semantic map method is to turn Table 1 into an informative map, which will reflect the regular cross-

**TABLE 1** Lexemes for tree/wood/forest in four languages

Lexical items			
Danish	French	German	Spanish
<i>træ</i>	<i>arbre</i>	<i>Baum</i>	<i>árbol</i>
	<i>bois</i>	<i>Holz</i>	<i>madera</i>
<i>skov</i>		<i>forêt</i>	<i>Wald</i>
	<i>selva</i>		

linguistic relations between the meanings of these language-specific lexical items. From Hjelmslev's relativism to the kind of universalism postulated by the semantic map model (see Section 3), there are just two steps, which should be taken in the following order.

## 2.1 | Identifying the meanings (nodes of the map)

The individual nodes (or vertices) of a semantic map are inferred from empirical evidence. Meaning identification is based on the *analytical primitive* principle (Cysouw, 2007, 2010a). According to this principle, a node N is an analytical primitive, if it cannot be subdivided into two (or more) meanings that are expressed by separate linguistic items in a given language. In practical terms, this means that a new node can be added to the map if and only if there is at least one language with a dedicated linguistic form for this node (Haspelmath, 2003; see further François, 2008). This principle therefore ensures that distinctive meanings will be as *linguistically* relevant as possible and will not just rely on linguists' idiosyncratic analyses. In Table 2, for instance, in the absence of Spanish, a distinction between the meaning WOOD (material) and FIREWOOD would not be justified. It is indeed the sole language in this table with a specific linguistic form for these two meanings, while Danish, French, and German have a single lexical item to express both meanings.

In accordance with the *analytical primitive* principle, five nodes can be identified based on the small language sample in Table 1. We use English as metalanguage and label these nodes: TREE, WOOD (material), FIREWOOD, FOREST (small), and FOREST (large). The semantic map method is neutral as regards the interpretation of these nodes or meanings. Some linguists see them as cognitively salient (Croft, 2001; but see Cristofaro, 2010, for a critique of this position), while others consider them to be merely comparative concepts (Haspelmath, 2010, 2016), specifically created by linguists for the purpose of comparing language-specific categorizations in the semantic domain (see Section 7).

**TABLE 2** Partitioning of the TREE–WOOD–FOREST semantic domain

		Lexical items			
		Danish	French	German	Spanish
ANALYTICAL PRIMITIVES	TREE	<i>træ</i>	<i>arbre</i>	<i>Baum</i>	<i>árbol</i>
	WOOD (mat.)		<i>bois</i>	<i>Holz</i>	<i>madera</i>
	FIREWOOD	<i>skov</i>		<i>Wald</i>	<i>leña</i>
	FOREST (small)		<i>forêt</i>		<i>bosque</i>
	FOREST (large)			<i>selva</i>	

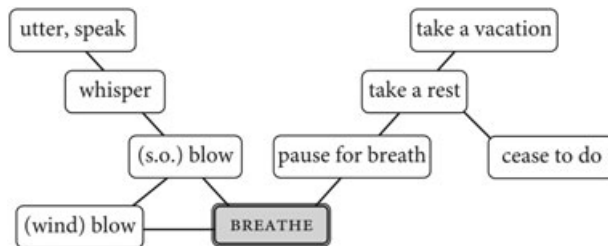
**TABLE 3** Lexical matrix for TREE/WOOD/FOREST in four languages

		Meanings				
		TREE	WOOD (mat.)	FIREWOOD	FOREST (small)	FOREST (large)
Danish	<i>træ</i>	√	√	√	-	-
	<i>skov</i>	-	-	-	√	√
French	<i>arbre</i>	√	-	-	-	-
	<i>bois</i>	-	√	√	√	(√)
	<i>forêt</i>	-	-	-	(√)	√
German	<i>Baum</i>	√	-	-	-	-
	<i>Holz</i>	-	√	√	-	-
	<i>Wald</i>	-	-	-	√	√

It is worth noticing that, when building a semantic map, both the onomasiological and the semasiological approaches can be used independently or combined (de Haan, 2010; Zwarts, 2010b, p. 124): with a *top-down* (onomasiological) approach, a given semantic/functional domain is investigated and the relevant linguistic expressions are listed (and subsequently structured) for each language; with a *bottom-up* (or *semasiological*) approach, language-specific grams, lexemes or constructions, and their multiple meanings are the starting point.

Most studies first proceed onomasiologically: they pick a particular domain, identify the core meanings in this domain, and search for the individual forms that express these meanings in different languages. In a second step, the semasiological dimension usually kicks in: one lists in a lexical matrix (Table 3) all the meanings attested for each form of the language sample. In such lexical matrices, if there are two or more √s in the same row, it means that the linguistic form is polysemous, while if there are two or more √ in the same column, the related linguistic forms are synonyms or translational equivalents.

Instead of picking one whole domain, other studies choose a single meaning as a pivot of the map. Taking as a point of departure the intra-linguistic onomasiological perspective, these studies first ask what are the words that express the meaning in question in a particular language. In the subsequent semasiological analysis, they list the different meanings of the relevant linguistic items in a language. The final step includes repeating this two-step process in the whole language sample chosen (see François, 2008; Georgakopoulos et al., 2016; Rice & Kabata, 2007). An illustration of the result of such an approach is provided in Figure 2.



**FIGURE 2** A (partial) semantic map for BREATHE combining both the onomasiological and the semasiological approach (François, 2008, p. 185)

This semantic map resembles a language-specific polysemy network, one of the differences being that the pivot (the notion BREATHE in Figure 2) is not similar to the prototypical meaning (see François, 2008, p. 181).

## 2.2 | Linking the meanings (edges of the graph)

In the semantic map model, the process of linking nodes follows a principal constraint known as the *connectivity hypothesis*: “any relevant language-specific and construction-specific category should map onto a CONNECTED REGION in conceptual space” (Croft, 2001, p. 96), “more precisely, a connected subgraph” (Croft & Poole, 2008, p. 4). As Andrason (2016, p. 2) puts it, the meanings

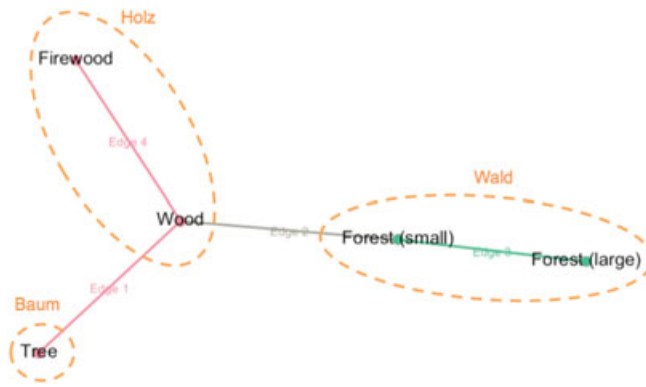
*are connected because they arise due to human cognitive mechanisms, being derived by means of metaphor, image-schema process, metonymy, analogy or abduction [... and o]n the other hand, they constitute a temporally sequential chain of predecessor and successors.*

As such, conceptual and historical factors support the connectivity hypothesis. In practical terms, this means that polysemous linguistic items are decisive when plotting a map. Indeed, they are the ones that will be mapped onto two (or more) nodes, and they indicate thereby which nodes should be connected: by virtue of the connectivity hypothesis, they must cover a connected region in the semantic map. Based on the data in Table 2, one can induce the following edges (Figure 3): the meanings TREE and WOOD can be connected, because they are expressed by a single word in Danish, the polysemous item *træ* (edge 1); the same applies to the meanings WOOD and FOREST (small), which can be linked because of the French polysemous item *bois* (edge 2), and to the meaning FOREST (small) and FOREST (large), because of the Danish and German lexemes *skov* and *Wald* (edge 3). The boundaries delimited by particular linguistic items in a language are conventionally represented by closed curved lines. For example, the boundaries of the German lexical items *Baum*, *Holz*, and *Wald* are shown in Figure 4. As expected, they do cover strictly connected regions of the map.

A second principle at work when plotting semantic maps is what we call the *economy principle*: given three meanings (Meaning<sub>1</sub>, Meaning<sub>2</sub>, Meaning<sub>3</sub>), if the linguistic items expressing Meaning<sub>1</sub> and Meaning<sub>3</sub> always express Meaning<sub>2</sub>, there is no need to draw an edge between Meaning<sub>1</sub> and Meaning<sub>3</sub> (the resulting map will be linear, Meaning<sub>1</sub>—Meaning<sub>2</sub>—Meaning<sub>3</sub>, and not triangular, with all the meanings connected). For example, even if Danish *træ* would allow us to directly connect the nodes FIREWOOD to TREE, and although French *bois* could lead



FIGURE 3 A semantic map inferred from the data in Table 2



**FIGURE 4** A semantic map inferred from the data of Table 2, with the German lexemes mapped onto the nodes

to linking FIREWOOD to FOREST (small), a single edge between FIREWOOD and WOOD (edge 4) is actually enough in order for the *connectivity hypothesis* to be respected. Then the only reason to draw an edge between the meanings FIREWOOD and TREE, or between FIREWOOD and FOREST (small), would be to identify a language in which these two meanings are expressed by a specific lexeme, which would crucially not express the meaning WOOD (that acts presently as an intermediate node between these two meanings).

The semantic map in Figure 3 was plotted in a strictly inductive fashion (which is called the “matrix-driven” approach in Zwarts, 2010a, pp. 378–379). In practice, however, one can observe “a combination of deductive semantic analysis and inductive generalizations on a sufficiently large sample of languages” (van der Auwera & Temürçü, 2006, p. 132). Some semantic map-like networks were even entirely developed following a deductive method (which is called the space-driven approach in Zwarts, 2010a, pp. 379–382): they are either based on extra-linguistic data (e.g., the organization of color chips into a color space according to physical features of hue, saturation, and brightness; e.g., Regier, Kay, & Khetarpal, 2007) or are the product of pre-empirical conceptual analysis (e.g., Lakoff, 1987, on the English preposition *over*). They can be considered as a good starting point for plotting actual semantic maps but should be tested against cross-linguistic data in order to assess the empirical validity of their claims regarding the organization of the semantic level.

### 3 | WHAT ARE THE ADVANTAGES OF THE SEMANTIC MAP MODEL?

The semantic map model is not a theory of grammar, but as Cysouw (2007, p. 228) phrased it, “a model of attested variation, which might [...] be the basis for the formulation of a theory.” It has several significant advantages that make it a useful tool to describe both “language universals and language-specific grammatical knowledge” (Croft, 2003, p. 133). In what follows, we synthesize its main advantages.

#### 3.1 | Advantages of semantic maps as a typological method

A first advantage is that it is neutral with respect to the monosemy/vagueness–polysemy–homonymy distinction (Haspelmath, 2003). A *monosemic* approach would consider the different

meanings of a form as being contextually driven (based on a vague or underspecified meaning); a *polysemic* account would recognize that different related meanings are associated with each lexical item; a *homonymic* position would argue that each meaning of a linguistic item on the map corresponds to a single form.<sup>3</sup> By not taking sides, the semantic map model gives a way out of the problems arising in adopting one of the stances. More specifically, its neutral perspective facilitates cross-linguistic comparison, an area in which the aforementioned approaches have little to offer. The very general meanings identified in monosemic analyses and the more sophisticated (but pertaining to language-specific grammars) networks constructed in studies that favor polysemic analyses, albeit both useful in some contexts, are not well suited for comparing languages (see Haspelmath, 2003, pp. 213–214, 230–232).

An additional advantage stemming from the neutral character of semantic maps is that they can be fruitfully used in various frameworks. Most scholars merely employ them as a tool (*tertium comparationis*) for studying cross-linguistic (as well as language-specific) patterns of polysemy, while remaining agnostic and refraining from any claim about their cognitive reality or universalism (Cysouw, 2007, p. 227). Other scholars, on the contrary, argue that the network of meanings can be envisioned as a universally valid organization of conceptual knowledge across languages, a “geography of the human mind,” as Croft (2001, p. 139) puts it. Designations such as “cognitive map” (Kortmann, 1997), “conceptual space” (Croft, 2001, 2003), or “mental map” (Anderson, 1986) are representative of this trend. Semantic maps are in this case understood similarly as the “networks” typical of cognitive grammar approaches (e.g., Langacker, 1988; Sandra & Rice, 1995). Yet other scholars are explicitly critical of the position that semantic maps give access to a universal arrangement of different conceptual situations in a speaker’s mental representation (Cristofaro, 2010; see also Janda, 2009). Despite the disagreements on how far the model can go, semantic maps constitute a suitable model for every approach mentioned.

Furthermore, the meanings or nodes of semantic maps can be of any kind, that is, “grammatical,” “lexical,” or “constructional” (see Section 6). Semantic maps can be used for any sort of structured semantic relationships. As a result, any area of the language can be investigated with a single tool, and there is no need to discriminate between the various kinds of meanings, the boundaries of which are not always clear-cut anyway.

Yet another advantage of semantic maps as a typological method is that they are at the same time *implicational* (Haspelmath, 1997a) and *falsifiable* (Cysouw et al., 2010, p. 1). This means that they articulate implicational hypotheses that are deemed to be universally valid as long as they are not falsified, that is, contradicted by new empirical evidence. For example, based on Figure 2, one can hypothesize that, if a language-specific lexical item expresses both the meaning TREE and the meaning FIREWOOD (like Danish *træ*), then it will necessarily express the meaning WOOD. If a given language turns out to have a single form expressing the meanings TREE and FIREWOOD, but not the meaning WOOD, then the map has to be emended<sup>4</sup> and new implicational universals can be formulated.

<sup>3</sup>We make a distinction here between homonymic interpretations of a map and the purposeful integration of homonyms in a single map, which is admittedly a problem since it generates uninformative maps, as discussed by van der Auwera & Temürçü (2006, p. 133) and van der Auwera, Kehayov, & Vittrant (2009, pp. 297).

<sup>4</sup>The map does not need to be revised, if (a) one is dealing with homonyms, or (b) it can be shown that this meaning was present in the language at some point in the past but has been taken over by another form (borrowed or not); see the discussion on Figures 10a–d below.



### 3.2 | Advantages of semantic maps as a semantic method

As shown in Section 2, semantic maps allow one to combine the *onomasiological* and *semasiological* perspectives, thus offering a semantically holistic view (see Lehmann, 2004; Geeraerts, 2010, p. 23; with Gast, 2009, pp. 212–213, specifically about semantic maps). The method proves directly useful both to answer the question of how languages express particular meanings or entire semantic fields (onomasiology) and to chart the different meanings of particular linguistic units in a given language (semasiology). In our example (Tables 1, 2, Figure 1), the onomasiological analysis reveals that the meaning WOOD is designated by the lexical items *træ* (Danish), *bois* (French), *Holz* (German), and *madera* (Spanish). Additionally, it gives intra-linguistic information, in that it indicates, for instance, that *bois* and *forêt* in French are near-synonyms for the meaning FOREST (onomasiological viewpoint). The semasiological analysis, on the other hand, shows that the lexical unit *træ* (Danish) covers three meanings. It also reveals that there are polysemic patterns recurring cross-linguistically, as indicated by the case of the Danish *skov* and the German *Wald* covering a similar region of the map. To sum up, with semantic maps, we are able to search for translational equivalents cross-linguistically and designations of a particular meaning intra-linguistically, on the one hand, and for regular and language-specific regular polysemy patterns (Cysouw, 2010b; Perrin, 2010), on the other hand (Table 4).

### 3.3 | Semantic maps and historical linguistics

Finally, semantic maps have proven to be an efficient tool in historical linguistics, and especially in grammaticalization studies (e.g., Narrog & van der Auwera, 2011). Synchronic semantic maps can indeed be interpreted diachronically, as they make prediction about the meanings to which a given form could extend, and a proper methodology has been elaborated for diachronic semantic maps, which explicitly visualize the attested pathways of evolution. This approach is discussed in Section 4.

## 4 | LINKING MEANINGS WITH SEMANTIC MAPS: TYPES OF RELATIONSHIPS, DIACHRONY, AND FREQUENCY

The semantic maps discussed in Sections 2 and 3 are classical semantic maps (also known as “traditional” in Malchukov, 2010; “first generation” in Sansò, 2010; “implicational” in Wälchli, 2010; or “connectivity maps” in van der Auwera, 2013). They usually take the form of two-dimensional graphs, with nodes (technically called “vertices”) connected by lines (technically called “edges”).

In the simplest form of classical semantic maps, the nodes are generally displayed with (or as) labels referring to a meaning, their precise position does not matter, and the length of the lines between nodes is irrelevant (Haspelmath, 2003, p. 216). The graph structure is the only aspect that really matters—formally speaking, classical semantic maps are undirected graphs

**TABLE 4** The semasiological and onomasiological features of semantic maps

	Cross-linguistic	Intra-linguistic
Onomasiology	Translational equivalents	Synonymy and near-synonymy
Semasiology	Regular polysemy patterns	Structured polysemy patterns

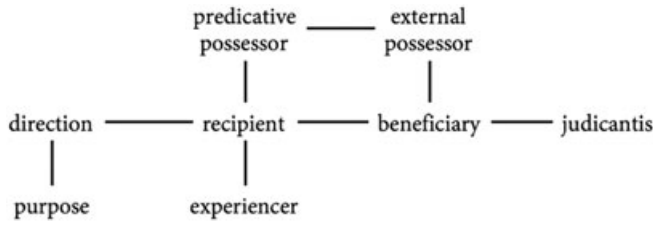


FIGURE 5 Semantic map of dative functions (adapted from Haspelmath, 2003, p. 213)

—which means that the similarity between two meanings depends on the number of intervening nodes (van der Auwera, p. 2013, 156). Thus, in Figure 1a, which we repeat here as Figure 5 for convenience, the distance between PURPOSE and EXPERIENCER is greater than the distance between PURPOSE and DIRECTION, because one has to pass two nodes to reach the former and none to reach the latter. The meanings PURPOSE and DIRECTION can thereby be inferred to be semantically closer than PURPOSE and EXPERIENCER. As stated above, the precise position of the node on the plane is not meaningful in this mode of representation. In Figure 5, for instance, the spatial distance between PURPOSE and EXPERIENCER is (more or less) the same as the one between PURPOSE and DIRECTION, but this only reflects an arbitrary positioning of the nodes and cannot be taken as evidence for proximity in meaning: the number of edges between nodes is the only thing that matters.

Several visualization techniques have been used in order to expand this basic type of representation and to capture graphically more information while remaining within the classical semantic maps model. In the literature, these techniques apply to three main kinds of information: (a) information about the types of relationships between the meanings, (b) diachronic information, and (c) information about the frequency of polysemy patterns.

### 4.1 | Types of relationships

In order to visualize different types of relationships between meanings, van der Auwera and Plungian (1998) represented meanings with elementary set-theoretical means: the inclusion of one oval into another indicates a hyper-/hyponymic relationship, while connecting two ovals with a line points to a metonymical (or metaphorical) link (van der Auwera, 2013, pp. 161–162).

In Figure 6, DEONTIC POSSIBILITY (e.g., “as far as I’m concerned, you may go to the party tonight”) is defined as a subtype (*hyponym*) of PARTICIPANT–EXTERNAL POSSIBILITY (e.g., “you may take the bus in front of the train station”), while PARTICIPANT–EXTERNAL POSSIBILITY and EPISTEMIC POSSIBILITY (e.g., “he may be at the office right now”) are seen as metonymically related. As observed by Zwarts (2010a, pp. 384–385), these types of semantic relationships could be represented as well by different types of lines. Figure 7 is a (visually less expressive) translation

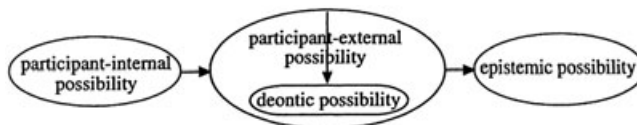
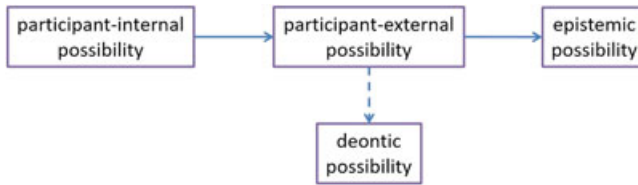


FIGURE 6 A mini-map of modal possibility (van der Auwera & Plungian, 1998, p. 87, Figure 1)



**FIGURE 7** A mini-map of modal possibility (adapted from Figure 6)

of Figure 6. Adhering to this alternative representational mode, a dashed line is used for the hyper-/hyponymic relationship, while a solid line is used for the metonymic (or metaphorical) links. Note that this visualization technique is not ideal for unbalanced semantic relations, like the one between a generic term and a more specific one, since one loses information about which node is the hypernym and which node is the hyponym.

**4.2 | Diachrony**

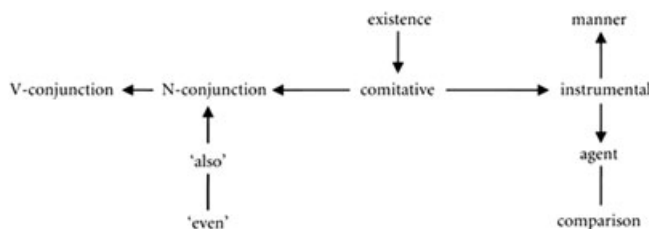
The example of Figure 6 displays another striking feature: the nodes are not connected by mere lines, but by arrows. This graphic device is used to integrate diachronic information about directionalities of change. Adding information about diachrony in a map is known as “dynamicizing” a map (Narrog & van der Auwera, 2011, pp. 323–327). Drawing from the terminology of graph theory, we define a dynamic semantic map (a *dysemap*) as a set of vertices connected by edges that are allocated a direction. These directed edges are called “arcs” and can represent different types of semantic shifts, such as “semantic generalization” or “specialization” in the case of hyper-/hyponymic relationships, or “semantic extension” when metaphorical and metonymical processes are involved (Luraghi, 2014; van der Auwera, 2013). Ideally, the dysemap would behave like a common directed graph (*digraph* in graph theory terminology), in that one single direction would be imposed on every edge (cf. Figure 8a–b), which is often the case for the semantic maps about grammaticalization pathways (that are largely unidirectional).

However, due to lack of data, it can happen that no directionality can be established between some vertices of a dysemap (in this respect, see the overlooked connections discussed in Narrog, 2010, p. 242, and Figure 9 below), or that, due to controversial directionalities (e.g., Narrog, 2010) or attested bi-directionalities (e.g., van der Auwera & Plungian, 1998, pp. 100, 111; Luraghi, 2001, p. 50; van der Auwera, Kehayov, & Vittrant, 2009, Maps 6 and 10), a double-headed arc connects a pair of vertices.

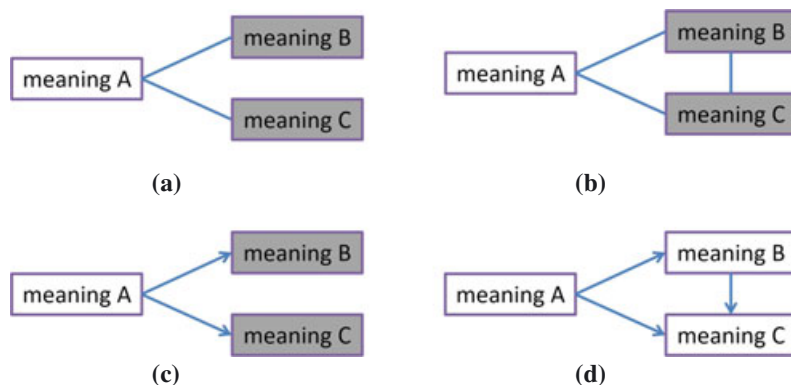
Even if only a small portion of semantic map research has tried to integrate the diachronic dimension so far (see Section 6), these efforts turn out to be crucial from a methodological point of view (van der Auwera, 2008; van der Auwera, 2013, pp. 164–167), since they allow one to explain exceptions to the connectivity hypothesis (Section 2). Let us consider an abstract



**FIGURE 8** (a) A simple dysemap (Narrog, 2010, p. 234). (b) A simple digraph (Balakrishnan & Ranganathan, 2012, p. 40)



**FIGURE 9** A semantic map for conjunction and related functions (Haspelmath, 2004, p. 21), with added directionalities (Narrog & van der Auwera, 2011, p. 326)



**FIGURE 10** (a) A simple semantic map. (b) A vacuous semantic map. (c) A simple dysemap. (d) An oriented dysemap

example in order to illustrate this point. In the hypothetical scenario of the synchronic semantic map of Figure 10a, in which meaning A is connected to both meanings B and C, imagine the case of a linguistic item expressing both MEANING B and MEANING C (shaded in Figure 10a), but not MEANING A. One would have to posit an edge between MEANING B and MEANING C, making the map vacuous (Figure 10b) and much less informative, since all the meanings are connected. A dysemap approach of the same meanings, however, will allow formulating the hypothesis that both MEANING B and MEANING C attested in synchrony for a linguistic item derive from an earlier meaning A (without having to draw an edge between those meanings, at least provisionally; cf. Figure 10c, d). Consequently, the strictly inductive, matrix-driven approach cannot be straightforwardly applied with the dysemas.

Another advantage of the dysemas is that, even if all the meanings are connected (Figure 10d), they allow generalizations that would not be possible with vacuous synchronic semantic map (like Figure 10b; cf. Narrog, 2010, pp. 234–235). Figure 10d, for example, illustrates the fact that MEANING C is a semantic extension of either MEANING A or B but makes the prediction that the opposite semantic shift is not possible.

### 4.3 | Frequency

Besides the representation of different types of semantic relationships and the visualization of dynamic links between nodes, classical semantic maps can also integrate information about the frequency of polysemy patterns. As stressed by Cysouw (2007, p. 232), in traditional semantic

maps, “the boundary between attested and unattested is given a very high prominence,” since the unique attestation of a polysemy pattern will be represented on the map exactly as a very common one, namely, with a simple edge between two nodes (see further Croft & Poole, 2008).

In order to address this issue, information about the frequency of polysemy patterns can be visualized in three different ways, using the length of the edges (called “proximity” in van der Auwera, 2013, pp. 156–157), the types of the edges, or the thickness of the edges.

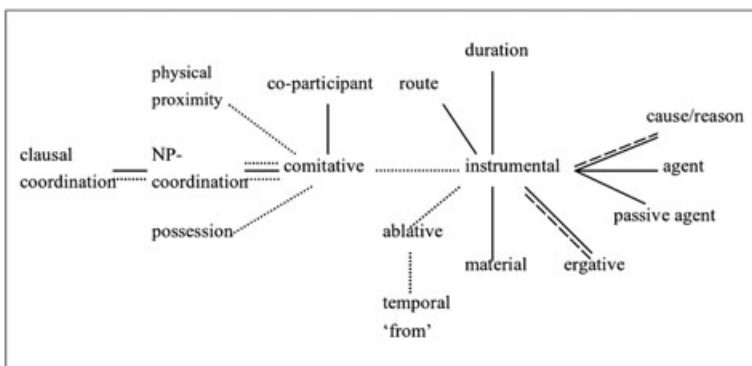
Figure 11 illustrates the length strategy: the difference in length of the edges between the nodes captures the cross-linguistic tendency for GOALS and PLACES to receive identical encoding, which is not so robust for PLACES and SOURCES (see Nikitina, 2009, pp. 1116–1117). Semantically, the semantic roles GOAL and PLACE will then be understood as more tied than PLACE and SOURCE.

In Figure 12, one observes different types of edges—solid lines, square dotted lines, round dotted lines, and long dashed lines—in order to represent different degrees of dependency of one meaning to another (Narrog & Ito, 2007, pp. 281–282). The solid lines, for example, indicate that a meaning depends on another one by more than 90%, with at least ten morphemes for which both meanings are available in a dataset of 200 languages; the square dotted lines, on the other hand, allow visualizing the dependency between three meanings (and not two), available in at least five morphemes (CLAUSAL COORDINATION–NP-COORDINATION–COMITATIVE is an example of such a dependency). As can be observed, including information about the frequency of different kinds of polysemy patterns leads to the multiplication of the number of edges between nodes.

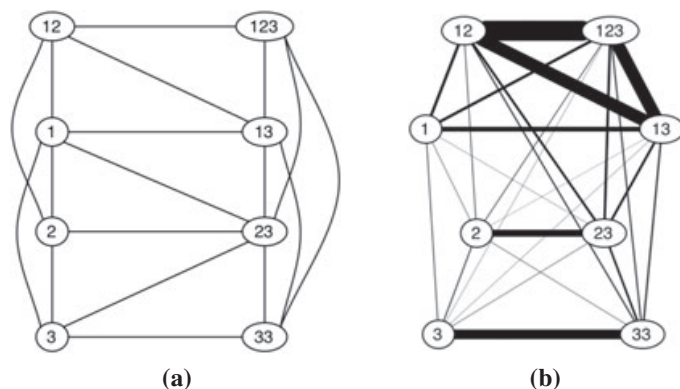
In his map of person marking, Cysouw (2007) provides a third solution for representing frequency, by using weighted edges whose thickness is proportional to the frequency of occurrence of the meaning pairs (see also Forker, 2016, p. 87).

Goal – Place — Source

**FIGURE 11** One-dimensional semantic map in which the length of the edges is meaningful (Nikitina, 2009, p. 1116)



**FIGURE 12** Visualizing different types of frequency in the semantic map of the Comitative-Instrumental domain (Narrog & Ito, 2007, p. 283)



**FIGURE 13** (a) A simple semantic map of person marking. (b) A weighted semantic map of person marking

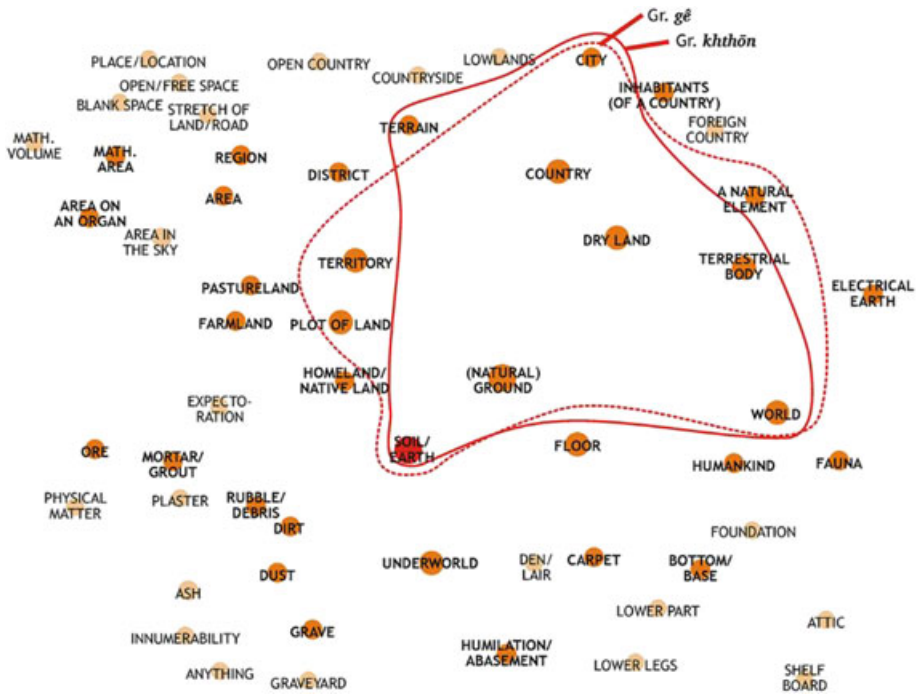
In the maps of Figure 13a,b, the numbers correspond to different primitives, which reflects the linguistic diversity of person marking in the languages of the world. The weighted edges in Figure 13b capture the frequencies of each polysemy pattern. The difference in thickness between the edge that connects node 1 (SPEAKER) to 12 (DUAL INCLUSIVE) and node 12 (DUAL INCLUSIVE) to 123 (PLURAL INCLUSIVE) represents the difference in frequency of colexification of the two pairs of primitives across languages (Cysouw, 2007, pp. 232–234). Not only is this kind of weighted classical semantic map much more informative than simple semantic maps, but it also allows one to simplify the map for the sake of generalizations, based on a principled criterion, namely, focusing on the more frequent polysemy patterns (see Malchukov, 2010, p. 177 about data reduction).

The comparison between Figure 13a,b further reveals that the two-dimensional semantic maps, which are almost unanimously preferred<sup>5</sup>—since they are easier to represent and read on paper (see Haspelmath, 2003, p. 218; Narrog & Ito, 2007, p. 273)—can be hard to interpret when nodes are densely interconnected and edges cross (technically called “non-planar graphs”). In this case, the readability of three-dimensional semantic maps such as Figure 13b is assuredly better.

Strongly connected maps, that is, maps in which some nodes are connected to many other nodes, can also be difficult to read and interpret. To avoid this state of affairs, which is a frequent and notable problem when studying the semasiology of a few lexical items in a given semantic field, the related meanings can also be visualized as neighboring meanings, albeit without connecting lines. In Georgakopoulos et al. (2016), visualizations of this type were possible using a semiautomatic process, which included both automatic and manual arrangements of meanings. In this case, the constraint is to arrange the meanings in such a way that it is possible to encircle contingent areas for all individual lexemes.

The map is valid as long as we are able to draw a closed curved line around all the meanings expressed by the lexemes, as is the case with Figure 14, which shows how Classical Greek lays down its own boundaries in the EARTH/SOIL domain. The same principle is applied in Tenser (2016, pp. 225–235) when studying the influence of Russian and Polish on the case representation system of Romani.

<sup>5</sup>See however the attempts to map in three dimensions of van der Auwera and Van Alsenoy (2013a, 2013b).



**FIGURE 14** The semantic space of EARTH/SOIL lexemes in Classical Greek (Georgakopoulos et al., 2016, p. 440)

## 5 | PROBLEMS WITH CLASSICAL SEMANTIC MAPS AND ALTERNATIVE MODELS

Starting with Cysouw (2001), the classical semantic map model discussed so far has been questioned and criticized for three main reasons.<sup>6</sup> First, “the precise predictions that can be formulated on the basis of [an] implicational map are unclear,” because it “predicts much more than is actually found” (Cysouw, 2001, pp. 609–610). To put it otherwise, the model is too strong for the data on which it is based: it overgenerates constellations, favoring high coverage over high accuracy (Cysouw, 2007, pp. 234–235; Croft & Poole, 2008, p. 6; Malchukov, 2010, p. 176). This point is easily illustrated based on the map of Figure 16a below. Theoretically, there are 105 different possibilities for mapping a linguistic form, whereas Haspelmath (2003, p. 76) states that only 39 different kinds of mapping are actually found in his dataset. Second, “as the amount of data increases, vacuous maps become more and more widespread since frequent, rare and exceptional patterns will all be represented on the map” (Malchukov, 2010, p. 176). Third, classical semantic maps could not be generated automatically at the time and were considered “not mathematically well-defined or computationally tractable, making it impossible to use with large and highly variable crosslinguistic datasets” (Croft & Poole, 2008, p. 1).

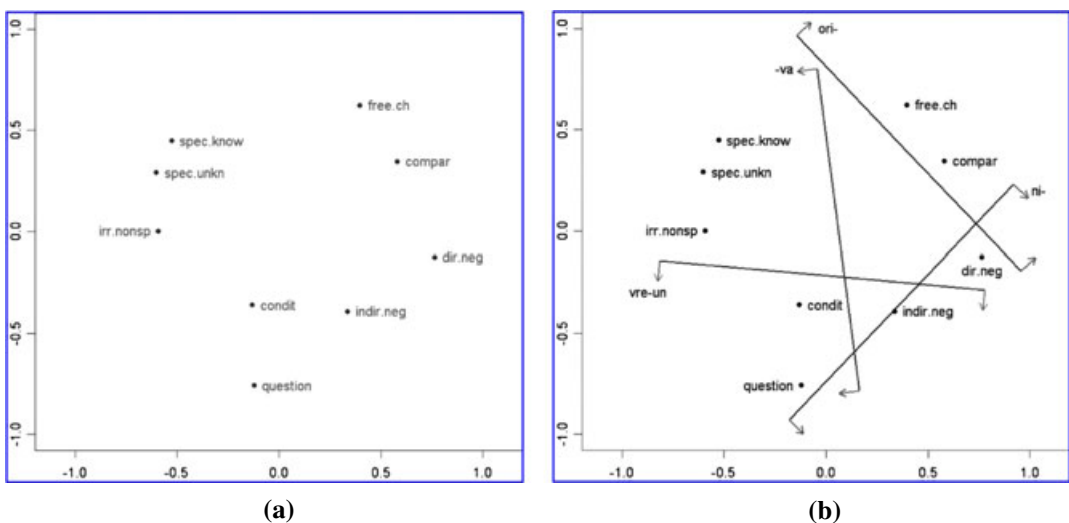
In order to cope with these issues, statistical scaling techniques—especially MDS—which are particularly efficient in dealing with big data, were introduced by several scholars as alternative

<sup>6</sup>We focus here exclusively on issues that have not been addressed in the previous sections. The issues connected to the visualization of frequency patterns, for instance, are dealt with in Section 4 (see also the discussion in Malchukov, 2010, pp. 176–177).

or complementary visualization methods. MDS is basically a means of visualizing spatially similarities and dissimilarities between pairs of items. To paraphrase Cysouw (2007, pp. 236, 241), the general idea behind the mathematical analysis is that the distance between two meanings in a two-dimensional Euclidian plane is iconic to the chance of co-occurrence of these meanings within a single linguistic expression (Schiffman, Reynolds, & Young, 1981; Groenen & van de Velden, 2005, p. 1280; Croft & Poole, 2008). Maps of this type are called, “similarity” (Malchukov, 2010, p. 176), “second generation” (Sansò, 2010), “statistical” (Wälchli, 2010), “probabilistic” (Wälchli & Cysouw, 2012), or “proximity” maps (van der Auwera, 2013).

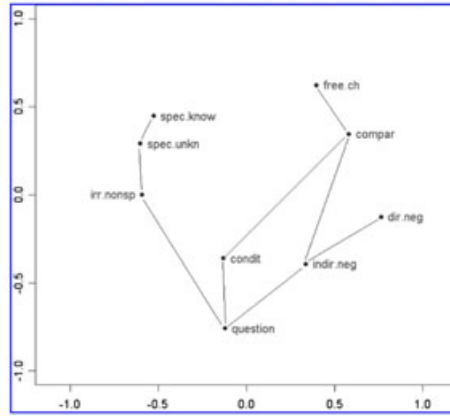
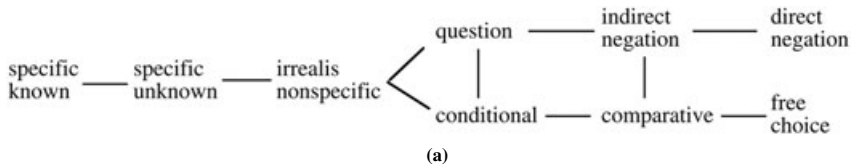
Figure 15a, taken from Croft and Poole (2008), exemplifies this visualization technique. It is based on Haspelmath’s (1997a) data used for the study of the semantics of indefinite pronouns. It tells us, among other things, that an indefinite expression occurs more frequently across languages with both the functions SPECIFIC, KNOWN TO THE SPEAKER (*spec.know*), and SPECIFIC, UNKNOWN TO THE SPEAKER (*spec.unkn*) than it does with both the function SPECIFIC, KNOWN TO THE SPEAKER (*spec.know*) and IRREALIS, NON SPECIFIC (*irr.nonsp*). In this case, the positioning of the various meanings on the two-dimensional plane is not the only product of MDS. An important aspect here is the addition of cutting lines, which correspond to the boundaries of the language-specific forms: as shown in Figure 15b for Romanian indefinite pronouns, these cutting lines fulfill the same function as the closed curved lines in the classical semantic map model (see Figures 1a, 4, and 14) showing which form expresses which function(s).

However different the classical semantic maps approach and the MDS procedure may seem, they can be thought of as compatible and complementary (Mauri, 2010; van der Auwera, 2008, 2013). In fact, they are able to represent the same structure of the conceptual space when visualizing the same data (Croft & Poole, 2008, p. 19). This point can be illustrated by comparing Figure 16a, namely, Haspelmath’s (1997a) original semantic map of the indefinite pronouns functions, with Figure 16b, the MDS analysis by Croft and Poole (2008) of the same data (cf. Figure 15a), with the superimposed graph structure of the classical semantic map. The curved horseshoe shape of the arrangement of the points in the two-dimensional MDS visualization is typical and explained by the fact that a single cutting line needs to be able to delimitate the



**FIGURE 15** (a) MDS analysis of Haspelmath’s (1997a) data on indefinite pronouns (Croft & Poole, 2008, p. 15). (b) Cutting lines for Romanian indefinite pronouns (Croft & Poole, 2008, p. 16)



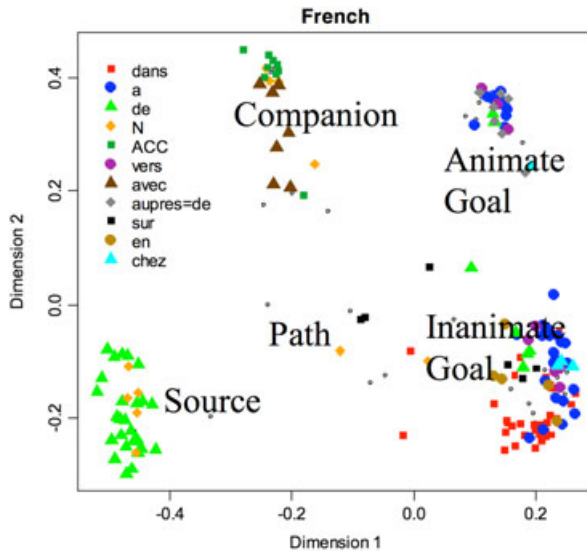


**FIGURE 16** (a) Haspelmath's (1997a, p. 4) original semantic map of the indefinite pronouns functions. (b) MDS analysis of Haspelmath's (1997a) data with the superimposed graph structure (Croft & Poole, 2008, p. 17, Figure 6)

language-specific categories in the Euclidian plane (cf. Figure 15b, with Croft & Poole, 2008, pp. 17–18).

However, the input for such proximity maps is often of a different nature. In most of the cases, proximity maps are not constructed on the basis of a lexical matrix with identified meanings that result from a preliminary analysis of the cross-linguistic material (like in the example just discussed). Rather, more frequently, they are compiled from responses to linguistic (e.g., Croft & Poole, 2008, pp. 22–31, who rely on Dahl's, 1985, database) or nonlinguistic materials (see Levinson et al., 2003, pp. 503–513; see also Majid, Boster, & Bowerman, 2008 for a similar semantic map-like approach, which uses correspondence analysis), or directly plotted based on parallel corpora (Wälchli, 2010, 2016; Wälchli & Cysouw, 2012). In this case, what is represented via MDS, that is, the analytical primitives, is the distribution of the actual coding means in context (and not meanings). This is a method for doing “typology without types” (Wälchli, 2010, p. 347; Wälchli & Cysouw, 2012, pp. 702–703).

Figure 17 nicely illustrates this method. The position of the points corresponds to the distribution of 190 motion event clauses from translations of the *Gospel of Mark* (153 languages from all continents) based on the (dis)similarity between the local phrase markers (adposition and/or case) used in each clause (i.e., in each specific context). The colors and shapes of the points, on the other hand, correspond to the mapping of the French coding means. Such a map must therefore be analyzed in two different ways. First, one has to explain the clustering of the points (for this particular example, the motion event clauses). The parameters are not given, but the result of the statistical analysis (Hamming distance as a distance measure, in this case), dimension 1 and dimension 2 need to be interpreted. Having studied the mapping of the local phrase markers on these points, Wälchli (2010, pp. 347–349) concludes that dimension 1 corresponds to SEMANTIC ROLES variation (as it distinguishes neatly SOURCES and GOALS), while dimension 2



**FIGURE 17** MDS visualization of the French local phrase markers in the *Gospel of Mark* (Wälchli, 2010, p. 348)

likely represents the combination of ANIMACY and LOCALIZATION (i.e., movement “to,” “unto,” “into”).<sup>7</sup> Figure 17 displays the result of this analysis with labels for the main clusters: COMPANION, (IN)ANIMATE GOAL, PATH, and SOURCE. In a second step, the mapping of the language-specific local phrase markers can be analyzed. To take a single example, one can observe in Figure 17 some uses of the French preposition *de* ▲ (“from”) in goal-oriented motion events. These outliers can be explained by the occurrence of this preposition in the valency pattern of *s’approcher de X* “to approach X” and in the compound preposition *de l’autre côté de* “at/to the other side of” (Wälchli, 2010, p. 147).

As stressed by Grossman and Polis (2012, p. 185) and exemplified by the discussion of Figure 17, the main difference between the classical semantic maps model and the distance-based representations is that the former is an *explanans*—being the result of cross-linguistic investigations and implying a semantic analysis that *precedes* the construction of the map—while the latter is an *explanandum* (cf. van der Auwera, 2008): the maps are plotted directly based on the data (which are constructed and not given, see Wälchli & Cysouw, 2012), and these represent the point of departure of the analysis. Consequently, distance-based maps are not implicational and cannot be used to constraint the data (Malchukov, 2010, p. 177).

It should be stressed that the MDS method has been criticized because it cannot take into account diachronic information, if available (Narrog, 2010; van der Auwera, 2008, 2013). For example, there is no way to infer any directionality from Figure 15a. The classical “connectivity” maps on the other hand predict that “a category can acquire a new function only if that function is adjacent on the semantic map to some function that the category already covers” (Haspelmath, 1997a, p. 129). The arrangement of the same meanings in Figure 16a indeed allows us to predict that, if an originally FREE CHOICE marker extends to also cover the QUESTION/CONDITIONAL

<sup>7</sup>The MDS visualization tries to show as much as possible of the actual distances but needs to convert the many dimensions of the dataset into a two-dimensional plane. Consequently, the dimensions can turn out to be difficult to interpret, and the emerging picture can turn out to be hard to read (cf. Cysouw, 2007, p. 237).

function, then that marker should first extend to cover the COMPARATIVE function. Thus, interpreted diachronically, classical semantic maps make predictions similar to the synchronic (implicational) maps. Up until recently (see Section 6), however, the statistical approach was the only way to handle large typological datasets and to generate automatically maps for studying cross-linguistic diversity.

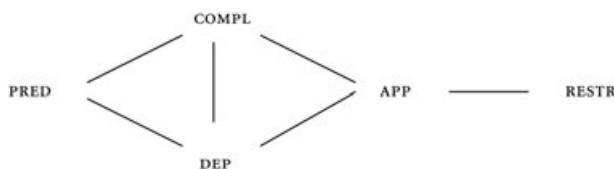
## 6 | SURVEY OF THE LITERATURE ON SEMANTIC MAPS

Any type of meaning can be integrated in semantic maps, such as the meanings of grammatical morphemes, of entire constructions, or of lexical items. From a methodological point of view, there is no need to distinguish among them, since the method can be used for any kind of structured semantic information.

Grammatical semantic maps cover a wide range of linguistic phenomena (cf. Cysouw, et al., 2010; Narrog & van der Auwera, 2011; van der Auwera & Temürçü, 2006, p. 132): *tense and aspect* (Anderson, 1982), *conditional markers* (Traugott, 1985), *reflexives and middles* (Kemmer, 1993), *indefinite pronouns* (Haspelmath, 1997a), *impersonal constructions* (Gast & van der Auwera, 2013; Malchukov & Ogawa, 2011; Siewierska & Papastathi, 2011; van der Auwera, Gast, & Vanderbiesen, 2012), *modality* (van der Auwera & Plungian, 1998; van der Auwera et al., 2009; Simon-Vandenberg & Aijmer, 2007, ch. 10; Boye, 2010), *temporal markers* (Haspelmath, 1997b), *encoding of core arguments* (Croft, 2001, pp. 134–147), *semantic roles* (Clancy, 2006; Georgakopoulos, 2014; Grossman & Polis, 2012; Hartmann, Haspelmath, & Cysouw, 2014; Haspelmath, 2003; Luján, 2010; Luraghi, 2001; Luraghi, 2014; Malchukov, 2010; Malchukov & Narrog, 2009; Mohammadirad & Rasekh-Mahand, 2017; Narrog & Ito, 2007; Rice & Kabata, 2007; Wälchli, 2010), *partitive constructions* (Koptjevskaja-Tamm, 2008), functions of generalized action verbs (Gil, 2017; Schultze-Berndt, 2008), *transfer of possession constructions* (Collins, 2015), *coordination* (Haspelmath, 2004, pp. 20–24; Mauri, 2010), *complementation* (Matras, 2004), *adversatives* (Malchukov, 2004), *intransitive predication* (Stassen, 1997), *secondary predication* (van der Auwera & Malchukov, 2005; Verkerk, 2009), *person-marking* (Cysouw, 2007), *imperative-hortatives* (van der Auwera, Dobrushina, & Goussev, 2003) *negative existentials* (Veselinova, 2013), *negative polarity items* (Hoekstra, 2014), *intensifying particles* (Forker, 2015), and *additives* (Forker, 2016).

As can be observed, many of the above grammatical semantic maps describe cross-linguistic polysemies of particular constructions rather than of isolated grammatical morphemes. Maps of this type allow one to capture which construction maps onto which category in a given language (see, e.g., Croft, 2001, ch. 2.4). Consider, for example, the semantic map in Figure 18 for depictive adjectival constructions proposed by van der Auwera and Malchukov (2005).

All the constructions visualized in the map belong to the same semantic domain. In compliance with the premises of the semantic map model, (a) the arrangement of the different types of predication in the graph—namely, PRED(icatives), DEP(ictives), COMPL(ementatives),



**FIGURE 18** Semantic map of depictive adjectivals (van der Auwera & Malchukov, 2005, p. 407)

APP(ositives), RESTR(ictives)—reflects the degree of (dis)similarity among these types; and (b) certain implicational hypotheses are possible. The map predicts that, if a language uses the same strategy for depictives and restrictives, then it will necessarily use the same strategy for appositives. An example of a language that uses the same adjectival strategy for all three types is English (ex. 1–3; van der Auwera & Malchukov, 2005).

1. Depictive: George left the party **angry**.
2. Appositive: My father, **angry** as always, left the party.
3. Restrictive (attributive): The **angry** young men left the party.

In fact, in English, all five types receive identical encoding (ex. 4–5).

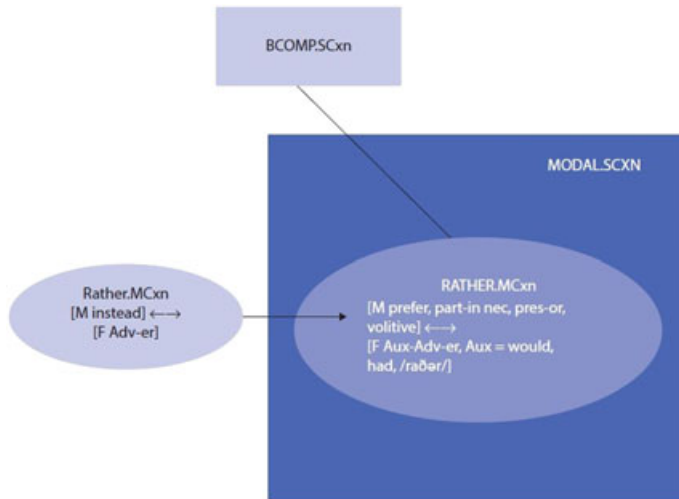
4. Complementative: I consider John **intelligent**.
5. Predicative: George was **angry**.

However, the different types of constructions yield many different permutations. In Russian, for instance, the instrumental forms of the adjective do not distinguish between depictives, predicatives, and complementatives, but they exclude appositives and restrictives (van der Auwera & Malchukov, 2005, p. 409).

In addition to grammatical and constructional maps, recent research has shown that the semantic map model can fruitfully be extended to lexical items. The starting point of this “lexical turn” can be traced back to François’ (2008) seminal paper, which, building on Haspelmath (2003), provides a blueprint for constructing lexical semantic maps (see Majid, Bowerman, van Staden, & Boster, 2007 for an early account; cf. Koch, 2001, for an approach similar to semantic maps). François uses semantic atoms or meanings of lexical items in context in order to analyze cross-linguistic patterns of colexification. Other studies that followed focused on polysemic patterns shared by diverse notions in different domains, such as quality expressions (Perrin, 2010; cf. Rakhilina, 2015; Ryzhova & Obiedkov, 2017), notions belonging to the motion domain (Wälchli & Cysouw, 2012) or to the domain of perception (Wälchli, 2016), the notion of emptiness (Rakhilina & Reznikova, 2014, 2016), temperature terms (Koptjevskaja-Tamm, 2015, p. 17; Liljegren & Haider, 2015, p. 469; Perrin, 2015), natural and spatial features (Georgakopoulos et al., 2016; Youn et al., 2016), and visual direction (Rakhilina, Vyrenkova, & Plungian, 2017).

It is fair to say that the different types of maps have not received equal attention in the literature. Rather, there is a strong bias towards studies describing cross-linguistic polysemies of grammatical morphemes and constructions. They have occupied a central role within the semantic maps tradition for at least two reasons. First, their study is often considered by linguists to be more interesting and prestigious than the study of the lexicon (Haspelmath, 2003, p. 211), and consequently, data about grammatical functions are more easily collected in the literature than data about polysemic lexical items. Second, the general tendency in the typology was to regard the lexicon as “exuberant and idiosyncratic” (François, 2008, p. 164). As a result, the lexical domain is not very well-studied, despite the fact that it has always been central for the arguments about cross-linguistic variation at the semantic level.

A common denominator to most of the studies listed above is their synchronic orientation. While it has been claimed that “the best synchronic semantic map is a diachronic one” (van der Auwera, 2008, p. 43; cf. Section 4 here, with Wälchli & Cysouw, 2012, pp. 703–705), the big bulk of research has been adopting a synchronic perspective, and the limited research that has added the diachronic dimension has focused almost exclusively on the grammatical domain



**FIGURE 19** Development of the micro-construction RATHER modeled as a MCM (Traugott, 2016, p. 120)

(Eckhoff, 2011; Guardamagna, 2016; Lichtenberk, 1991; Luján, 2010; Luraghi, 2014; Narrog, 2010; van der Auwera & Plungian, 1998). For lexical typology, semantic maps have been conceptualized explicitly as “a strictly synchronous device,” a stance justified by the complexity of the historical relations between lexical meanings (Rakhilina & Reznikova, 2016, p. 113; but see Viberg’s, 1984, modality hierarchy, which can be seen as a forerunner of lexical diachronic semantic maps). On the other hand, one can notice that the scope of constructional maps has been expanded in order to include the diachronic dimension (Fried, 2007, 2009; Traugott, 2016). Traugott (2016) shows how semantic maps can be used to inform diachronic constructional analyses. For these diachronic constructional maps to work, she argues that two levels are needed: a macro-level, which “represent[s] relationships between abstract, conceptual schemas linked to underspecified form” and a micro-level, “which models relationships among specific micro-constructions” (Traugott, 2016; see also Croft, 2001). These two levels correspond to two kinds of maps each operating at a different level of abstraction: the schema-construction maps and the micro-construction maps, respectively. In incorporating directionality of change, each type makes different generalizations: the former captures tendencies, and the latter language-specific paths of change (Traugott, 2016). Figure 19 presents a language-specific path of change in English modals of comparison. Through constructionalization *rather* extends from the uses “instead” and “sooner” (the [F Adv-er] on the map) to the modal use “d rather” [F Aux-Adv-er] (a development from non-volitive to volitive; cf. Narrog, 2012). The figure also captures the association of the RATHER micro-construction with two larger schema-constructions, the modal schema construction (MODAL.SCxn) and the biclausal comparative schema construction (BCOMP.SCxn), the latter of which is placed outside the modal domain.

## 7 | ISSUES, CHALLENGES, AND AVENUES FOR FUTURE RESEARCH

The great variety of linguistic domains to which the classical semantic map model has been applied highlights its efficiency in capturing regular patterns of semantic structure and cross-linguistic similarities of form-meaning correspondence. In this concluding section, we point

out some pending issues, challenges, and promising avenues for future research as regards (1) data collection, (2) the connectivity hypothesis, (3) automatic plotting, and (4) visualization techniques.

## 7.1 | Data collection

One major issue for the semantic map model, which is a recurrent concern in language typology as a whole, is the choice of a good language sample that will allow for valid cross-linguistic generalizations. Haspelmath (2003, p. 217) argues that a dozen of genealogically unrelated languages usually suffice to arrive at a certain degree of generalization. However, restricting typological research to only a few languages could result in overlooking interesting (even if infrequent) connections between meanings (Narrog & Ito, 2007, p. 276) or in missing linguistic or culture associations that are specific to geographical regions or areas. Narrog and Ito (2007, p. 276) suggest that the greater the size of the language sample, the greater the likelihood that the map will be accurate and capture (statistical) universals.<sup>8</sup> One important future area of research for the semantic map method would then be to construct and test various areally and genealogically stratified samples (on the language sampling method, see Rijkhoff & Bakker, 1998; Miestamo, Bakker, & Arppe, 2016, among others; cf. Bickel, 2015 for a caveat on representative samples). One question that will necessarily arise is whether lexical semantic maps should follow the same principles as grammatical semantic maps. In this respect, Rakhilina and Reznikova (2016, pp. 101–102) highlight the fact that some of the restrictions of grammatical typology do not apply to lexical typology. They claim, for example, that related languages can provide reliable information just as genealogically diverse ones do. Furthermore, despite the increasing availability of resources (such as the *Database of Cross-Linguistic Colexifications* [<http://clics.lingpy.org>], see List, Mayer, Terhalle, & Urban, 2014), the primary material for lexico-typological studies is not always sufficient, a factor that may impede large-scale studies. This is one of the main reasons why the number of languages of a typical lexico-typological study ranges from 10 to 50 (see Koptjevskaja-Tamm, Rakhilina, & Vanhove, 2015, p. 436; cf. Wälchli, 2010; Wälchli & Cysouw, 2012; Östling, 2016, which relied on larger samples, thanks to the availability of resources, viz. massively parallel texts).

Besides the quantity of data, the accuracy of a semantic map also depends heavily on the quality of the collected cross-linguistic material, which is best ensured by identifying comparable phenomena across languages. As to what counts as meaning, comparability is reached if the same definition is used, a definition that should ideally be purely descriptive and theory-neutral (see François, 2008, p. 170; Juvonen & Koptjevskaja-Tamm, 2016; Koptjevskaja-Tamm, 2016, p. 5). In this respect, the meanings of a map can be seen as comparative concepts (Haspelmath, 2010; see the special issue of *Linguistic Typology* 20/2 [2016] devoted to this topic), which have to be universally applicable and can be defined based on universal conceptual–semantic concepts, general formal concepts, and other comparative concepts.<sup>9</sup>

Yet two questions remain to be explored more thoroughly as regards data quality: on the one hand, the level of granularity of the meanings integrated in a semantic map, and on the other hand, the mapping of language-specific forms onto these meanings.

---

<sup>8</sup>In order to construct their map for the COMITATIVE-INSTRUMENTAL area, they relied on a sample of 200 languages.

<sup>9</sup>For example, a definition of a “future tense” as “[...] a grammatical marker associated with the verb that has future time reference as one prominent meaning” is based on the conceptual–semantic concept “future time reference” and the comparative concepts “verb” and “grammatical marker” (Haspelmath, 2010, p. 671).

The construction of a semantic map is indeed affected by decisions on the degree of resolution of the semantic distinctions (see Wälchli, 2010, p. 335). A map of higher resolution means that the analytical primitives used as a basis for plotting it are fine-grained, which leads to more detailed and accurate maps.<sup>10</sup> A map of lower resolution helps unravel general tendencies but will probably fail to capture more infrequent patterns (which however is not always considered as a problem; see François, 2008, pp. 163–164). While it is desirable to combine large cross-linguistic databases and a meticulous semasiological analysis of the selected linguistic items, thus obtaining a higher resolution, this is difficult to put into practice. Furthermore, despite some suggestions for visualizing hyper- and hyponymic relationships (see the discussion of Figure 6 above), the systematic integration of meanings of different degrees of generality within a single semantic map is still to be investigated.

As regards the mapping of language-specific forms onto the map, a recurring challenge for the method is that it often attributes meanings to the grammatical or lexical items themselves, despite the fact that we are usually dealing with contextual meanings that are only available for this form in specific constructions (cf. Grossman & Polis, 2012, p. 197). As Andrason (2016, p. 7) puts it, “[a] form that is represented by means of semantic maps is typically studied in isolation from the language in which it exists and in which it has been developing. (...) The lack of information concerning environmental factors is particularly suspicious (...)” A solution for integrating information about the construction-specific meanings of the forms that are mapped is yet to be found.

## 7.2 | The connectivity hypothesis

Another pending issue for the semantic map approach is how to account for violations of the connectivity hypothesis (Section 2). These violations can result from three main types of phenomena (e.g., van der Auwera, 2013, pp. 161–162): homonymy, diachrony, and language contact situations.

- Homonyms do not have to cover a connected region of a semantic map. Formal identity does not lead to semantic connectivity in cases such as *lie*<sub>1</sub> “speak falsely” and *lie*<sub>2</sub> “be positioned horizontally” (van der Auwera, 2013).
- As discussed in Section 4.2, dynamicized semantic maps, given their capacity to integrate the diachronic dimension, make it possible to explain the lack of connectedness between the meanings of a given linguistic forms in synchrony if (and only if) these meanings derive from a common “ancestor,” namely, a meaning previously expressed by the same form.
- In language contact situations, two types of exceptions to the connectivity hypothesis have been noticed in the literature. First, several scholars observed that areal factors possibly lead to the extension of the meaning of a linguistic form in a given language based on the meaning of a similar expression in a (prestigious) neighboring language (e.g., van der Auwera et al., 2009). This phenomenon, known as “polysemy copying,” has been studied within the classical semantic map method and described with the labels “semantic map harmony” (Tenser, 2008, 2016; see also Matras, 2009, pp. 263–264) and “semantic map assimilation” (Gast & van der Auwera, 2012). Second, in a study about adpositions borrowing between Greek and Coptic, Grossman and Polis (2017) showed that the

<sup>10</sup>See, in this respect, Wälchli and Cysouw’s (2012, p. 680) criticism: “[i]n implicational maps there are a small number of idealized functions that do not take into account the large amount of domain internal diversity of general abstract labels.”

polysemy network of the adpositions in the donor language is not borrowed as a whole; rather, only some of its meanings are borrowed, which are not necessarily connected on the map.

The clear identification of such exceptions is crucial for the semantic map method, as they directly bear on the automatic inference of semantic maps based on polysemy matrices (see Table 3).

### 7.3 | Automatic plotting

As already noted by Narrog and Ito (2007, p. 280), “ideally (...) it should be possible to generate semantic maps automatically on the basis of a given set of data.” Indeed, it is practically impossible to handle large-scale cross-linguistic datasets manually. However, as noted by Croft and Poole (2008, p. 7), it was at the time “not clear whether the semantic map model can be automated in a computationally tractable algorithm.” Finding the minimum number of links between nodes for a set of cross-linguistic data is akin to the “traveling salesman problem,” which is known to be NP-hard.<sup>11</sup> This potential intractability was considered to be a significant problem for the use of graph-based semantic maps in typology and led to the use of MDS (and similar techniques) for representing similarity between nodes (Section 5).

This state of affairs recently changed, when Regier, Khetarpal, and Majid (2013) showed that the semantic map inference problem is “formally identical to another problem that superficially appears unrelated: inferring a social network from outbreaks of disease in a population” (Regier et al., 2013, p. 91). This similar inference problem was shown to be indeed computationally intractable, but it was found that “an efficient algorithm exists that approximates the optimal solution nearly as well as is theoretically possible” (Angluin, Aspnes, & Reyzin, 2010). Having tested the algorithm on the cross-linguistic data of Haspelmath (1997a) and Levinson et al. (2003), Regier et al. (2013) concluded that the approximations produced by the algorithm are of high quality, which means that they produce equal or better results than the manually plotted maps. Hence, the graph structure of classical semantic maps can be quite straightforwardly inferred using such an algorithm.

However, very many questions remain to be explored in this highly promising domain. For instance, the algorithm of Regier et al. (2013) produces unweighted and undirected graphs: the automatic addition of weighted edges based on the cross-linguistic frequency of polysemy patterns and the inference of oriented edges based on diachronic information is shown to be both straightforward and highly informative in Georgakopoulos and Polis (2017). Besides, the problem of network inference is a very active research area (especially in biology, where network inference is used for uncovering causal relationships between genotype and phenotype), and the number of available algorithms has grown tremendously during the last decades (e.g., Siegenthaler & Gunawan, 2014). Such algorithms should be tested on large-scale cross-linguistic data in order to evaluate their efficiency in plotting informative maps.

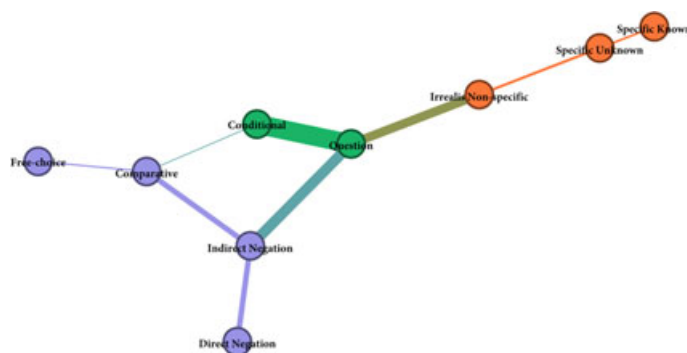
---

<sup>11</sup>NP-Hard stands for “Non-deterministic Polynomial-time Hard” problems, which refers to problems that are at least as hard as the hardest problem in NP.





**FIGURE 20** Visualizing diachrony and frequency in a single map (van der Auwera, 2013, p. 165, Figure 14)



**FIGURE 21** Visualization of Haspelmath's (1997a) data in Gephi (<https://gephi.org>) with the Force Atlas algorithm (including weighted edges and modularity analysis)

## 7.4 | Visualization techniques

As observed in Section 4, different kinds of linguistic information can be visually combined within a single semantic map. Figure 6 illustrated the fact that the types of semantic relationships between the nodes and diachronic data can be represented in the same map (see also van der Auwera, 2008; van der Auwera et al., 2009). Examples of the combination of diachronic and frequency data are not forthcoming. An abstract example is provided by van der Auwera (2008; see Figure 20), but empirical studies combining these two dimensions are virtually missing.<sup>12</sup>

Another visualization possibility that has not yet been explored is the simultaneous visualization of the type and of the frequency of polysemy patterns (let alone about other kinds of data). This is a promising avenue for future research. In the case, for example, in which one systematically assigns to the edges of a semantic map different flags referring to semantic relations such as metaphor, metonymy, and so on, a more thorough picture of the semantic domain(s) in question shall visually emerge. This could help us determine that some metaphors are more universal than others and that some are more culture-sensitive (cf. the COGNITION IS PERCEPTION primary metaphor vs. the UNDERSTANDING IS SEEING and UNDERSTANDING IS HEARING cultural-sensitive metaphors; see Evans & Wilkins, 2000; Ibarretxe-Antuñano, 2013; Sweetser, 1990; Vanhove, 2008).

Finally, graph visualization platforms have not been used for exploring classical semantic maps. These powerful tools, with many built-in statistical methods, reveal much information otherwise “hidden” in the network. The visual quality and quantity of information conveyed by Figure 21 can be conveniently compared to that of Figure 16a,b, which are based on the same dataset<sup>13</sup> (from Haspelmath, 1997a). The map of Figure 21 was plotted automatically using a modified version of Regier's et al. (2013) algorithm (that adds weights to the inferred edges)

<sup>12</sup>Note that the frequency accounted for in the “waves and streams” model introduced by Andrason (2016) is not the general frequency of polysemy patterns (like in Figures 13b and 21), but the frequency of individual forms or constructions (which would have to be integrated at the level of the mapping of specific forms onto the map in the semantic map model).

<sup>13</sup>We are most grateful to Martin Haspelmath for giving us access to his dataset and allowing us to make use of it.

and is visualized using the Force Atlas algorithm of Gephi (<https://gephi.org>) and modularity analysis (that measures the strength of division of a network into modules; for a detailed explanation, see Georgakopoulos & Polis, 2017). This basic example should suffice to show that visualization techniques and actual semantic analysis will be inseparable in the future of the semantic map model (Malchukov, 2010, p. 177).

## ACKNOWLEDGEMENT

The authors thank Eitan Grossman (Hebrew University of Jerusalem), two anonymous reviewers, and the editors for their insightful comments on this paper. Thanasis Georgakopoulos would like to acknowledge the support of the Marie Curie BeIPD postdoctoral grant from the European Commission. Stéphane Polis acknowledges the support of the National Fund for Scientific Research (Belgium).

## ORCID

Thanasis Georgakopoulos  <http://orcid.org/0000-0002-5632-4202>

Stéphane Polis  <http://orcid.org/0000-0003-0350-9753>

## REFERENCES

- Anderson, L. B. (1982). The 'perfect' as a universal and as a language-particular category. In P. J. Hopper (Ed.), *Tense-aspect: Between semantics & pragmatics, typological studies in language* (pp. 227–264). Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.1.16and>
- Anderson, L. B. (1986). Evidentials, paths of change, and mental maps: typologically regular asymmetries. In W. Chafe, & J. Nichols (Eds.), *Evidentiality: The linguistic encoding of epistemology* (pp. 273–312). Norwood: Ablex.
- Andrason, A. (2016). From vectors to waves and streams: An alternative approach to semantic maps. *Stellenbosch Papers in Linguistics*, 45, 1–29. <https://doi.org/10.5774/45-0-211>.
- Angluin, D., Aspnes, J., & Reyzin, L. (2010). Inferring social networks from outbreaks. In M. Hutter, F. Stephan, V. Vovk, & T. Zeugmann (Eds.), *Algorithmic learning theory 2010* (pp. 104–118). Berlin: Springer. [https://doi.org/10.1007/978-3-642-16108-7\\_12](https://doi.org/10.1007/978-3-642-16108-7_12)
- Ariel, M. (2008). *Pragmatics and grammar*. Cambridge: Cambridge University Press.
- Balakrishnan, R., & Ranganathan, K. (2012). *A textbook of graph theory*. Berlin: Springer.
- Bickel, B. (2015). Distributional typology: Statistical inquiries into the dynamics of linguistic diversity. In B. Heine, & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (2nd ed.) (pp. 901–923). Oxford: Oxford University Press.
- Boye, K. (2010). Semantic maps and the identification of cross-linguistic generic categories: Evidentiality and its relation to epistemic modality. *Linguistic Discovery*, 8(1), 4–22. <https://doi.org/10.1349/psl.1537-0852.a.344>
- Clancy, J. S. (2006). The topology of Slavic case: Semantic maps and multidimensional scaling. *Glossos*, 7, 1–28 Retrieved from <http://seelrc.org/glossos/>
- Collins, J. (2015). 'Give' and semantic maps. In B. Nolan, G. Rawoens, & E. Diedrichsen (Eds.), *Causation, permission, and transfer. Argument realisation in GET, TAKE, PUT, GIVE and LET verbs* (pp. 129–146). Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.167.04col>
- Cristofaro, S. (2010). Semantic maps and mental representation. *Linguistic Discovery*, 8(1), 35–52. <https://doi.org/10.1349/psl.1537-0852.a.345>

- Croft, W. (2001). *Radical construction grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, W. (2003). *Typology and universals* (2nd ed.). Cambridge: Cambridge University Press.
- Croft, W., & Poole, K. T. (2008). Inferring universals from grammatical variation: Multidimensional scaling for typological analysis. *Theoretical Linguistics*, 34(1), 1–37. <https://doi.org/10.1515/thli.2008.001>
- Cysouw, M. (2001). Martin Haspelmath, indefinite pronouns (Oxford studies in typology and linguistic theory). Oxford: Clarendon press, 1997. Pp. xvi 364. *Journal of Linguistics*, 37(3), 607–612. <https://doi.org/10.1017/s0022226701231351>
- Cysouw, M. (2007). Building semantic maps: The case of person marking. In B. Wälchli, & M. Miestamo (Eds.), *New challenges in typology* (pp. 225–248). Berlin/New York: Mouton de Gruyter.
- Cysouw, M. (2010a). Semantic maps as metrics on meanings. *Linguistic Discovery*, 8(1), 70–95. <https://doi.org/10.1349/ps1.1537-0852.a.346>
- Cysouw, M. (2010b). Commentary on Perrin—Drawing networks from recurrent polysemies. *Linguistic Discovery*, 8(1). <https://doi.org/10.1349/ps1.1537-0852.a.377>
- Cysouw, M., Haspelmath, M., & Malchukov, A. L. (2010). Introduction to the special issue “semantic maps: Methods and applications”. *Linguistic Discovery*, 8(1), 1–3. <https://doi.org/10.1349/ps1.1537-0852.a.358>
- Dahl, Ö. (1985). *Tense and aspect systems*. Oxford: Basil Blackwell.
- Eckhoff, H. M. (2011). *Old Russian possessive constructions: A construction grammar approach*. Berlin: De Gruyter Mouton.
- Evans, N., & Wilkins, D. (2000). In the mind's ear: The semantic extension of perception verbs in Australian languages. *Language*, 76, 546–592. <https://doi.org/10.2307/417135>
- Forker, D. (2015). Towards a semantic map for intensifying particles: Evidence from Avar. *STUF - Language Typology and Universals*, 68(4), 485–513. <https://doi.org/10.1515/stuf-2015-0021>
- Forker, D. (2016). Toward a typology for additive markers. *Lingua*, 180, 69–100. <https://doi.org/10.1016/j.lingua.2016.03.008>
- François, A. (2008). Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In M. Vanhove (Ed.), *From polysemy to semantic change. Towards a typology of lexical semantic associations* (pp. 163–215). John Benjamins: Amsterdam/Philadelphia.
- Fried, M. (2007). Constructing grammatical meaning: Isomorphism and polysemy in Czech reflexivization. *Studies in Language*, 41, 721–764. <https://doi.org/10.1075/sl.31.4.02fri>
- Fried, M. (2009). Plain vs. situated possession in a network of grammatical constructions. In W. McGregor (Ed.), *Expressions of possession* (pp. 213–248). Berlin: De Gruyter.
- Gast, V. (2009). A contribution to ‘two-dimensional’ language description: the typological database of intensifiers and reflexives. In M. Everaert, S. Musgrave, & A. Dimitriadis (Eds.), *The use of databases in cross-linguistic studies* (pp. 209–234). Berlin/New York: Mouton de Gruyter.
- Gast, V., & van der Auwera, J. (2012). What is ‘contact-induced grammaticalization’? Examples from Mayan and Mixe-Zoquean languages. In B. Wiemer, B. Wälchli, & B. Hansen (Eds.), *Grammatical replication and borrowability in language contact* (pp. 381–426). Berlin: de Gruyter Mouton.
- Gast, V., & van der Auwera, J. (2013). Towards a distributional typology of human impersonal pronouns, based on data from European languages. In D. Bakker, & M. Haspelmath (Eds.), *Languages across boundaries. Studies in memory of Anna Siewierska* (pp. 119–158). Berlin: de Gruyter Mouton.
- Geeraerts, D. (2010). *Theories of lexical semantics*. Oxford: Oxford University Press.
- Georgakopoulos, T. (2014). On the encoding of allative and recipient in the Greek diachrony. In S. Kutscher & D. A. Werning (Eds.), *On ancient grammars of space: Linguistic research on the expression of spatial relations and motion in ancient languages* (pp. 45–66). Berlin/Boston: De Gruyter.
- Georgakopoulos, T., & Polis, S. (2017). Inferring weighed and dynamic semantic maps, and taking advantages of graph visualization platforms: A couple of steps forward with the classical model. Unpublished manuscript.

- Georgakopoulos, T., Werning, A. D., Hartlieb, J., Kitazumi, T., van de Peut, E. L., Sundermeyer, A., & Chantrain, G. (2016). The meaning of ancient words for 'earth'. An exercise in visualizing colexification on a semantic map. *eTopoi. Journal for Ancient Studies*, 6, 1–36.
- Gil, D. (2017). Roon ve, DO/GIVE Coexpression and language contact in Northwest New Guinea. *NUSA*, 62, 41–100.
- Groenen, P. J., & Van de Velden, M. (2005). Multidimensional scaling. In B. S. Everitt, & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (Vol. II) (pp. 1280–1289). New York: John Wiley & Sons Ltd. <https://doi.org/10.1002/0470013192.bsa415>
- Grossman, E., & Polis, S. (2012). Navigating polyfunctionality in the lexicon. Semantic maps and Ancient Egyptian lexical semantics. In E. Grossman, S. Polis, & J. Winand (Eds.), *Lexical Semantics in Ancient Egyptian, Lingua Aegyptia Studia Monographica* 9 (pp. 175–225). Hamburg: Widmaier Verlag.
- Grossman, E., & Polis, S. (2017). Polysemy networks in language contact: The borrowing of the Greek-origin preposition κατά/κατα in Coptic. In E. Grossman, P. Dils, T. S. Richter, & W. Schenkel (Eds.), *Greek influence on Egyptian-Coptic: Contact-induced change in an Ancient African language* (pp. 335–367). Hamburg: Widmaier Verlag.
- Guardamagna, C. (2016). A diachronic semantic map for the Latin preposition secundum. *Journal of Latin Linguistics*, 15(2), 233–277.
- Haan, F. D. (2010). Building a semantic map: Top-down versus bottom-up approaches. *Linguistic Discovery*, 8(1), 102–117. <https://doi.org/10.1349/ps1.1537-0852.a.347>
- Hartmann, I., Haspelmath, M., & Cysouw, M. (2014). Identifying semantic role clusters and alignment types via microrole coexpression tendencies. *Studies in Language*, 38(3), 463–484. <https://doi.org/10.1075/sl.38.3.02har>
- Haspelmath, M. (1997a). *Indefinite pronouns*. Oxford: Oxford University Press.
- Haspelmath, M. (1997b). *From space to time: Temporal adverbials in the world's languages*. München: Lincom Europa.
- Haspelmath, M. (2003). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In M. Tomasello (Ed.), *The new psychology of language* (Vol. 2) (pp. 211–243). New York: Erlbaum.
- Haspelmath, M. (2004). *Coordinating constructions*. Amsterdam: John Benjamins.
- Haspelmath, M. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language*, 86(3), 663–687. <https://doi.org/10.1353/lan.2010.0021>
- Haspelmath, M. (2016). The challenge of making language description and comparison mutually beneficial. *Linguistic Typology*, 20(2), 299–303. <https://doi.org/10.1515/lingty-2016-0008>
- Hjelmslev, L. (1961). *Prolegomena to a theory of language*. Madison: The University of Wisconsin Press. English translation of *Omkring sprogteoriens grundlæggelse*, Copenhagen : Ejnar Munksgaard, 1943.
- Hoekstra, E. (2014). Semantic maps and negative polarity. In J. Hoeksema, & D. Gilbers (Eds.), *Black book. A Festschrift in honor of Frans Zwarts* (pp. 146–164). Groningen: University of Groningen.
- Ibarretxe-Antuñano, I. (2013). The relationship between conceptual metaphor and culture. *Intercultural Pragmatics*, 10(2), 315–339. <https://doi.org/10.1515/ip-2013-0014>
- Janda, L. (2009). What is the role of semantic maps in cognitive linguistics? In P. Stalmaszczyk, & W. Oleksy (Eds.), *Festschrift for Barbara Lewandowska-Tomaszczyk* (pp. 105–124). Frankfurt/Main: Peter Lang.
- Juvonen, P., & Koptjevskaja-Tamm, M. (Eds.) (2016). *The lexical typology of semantic shifts*. Berlin: De Gruyter.
- Kemmer, S. (1993). *The middle voice*. Amsterdam: Benjamins.
- Koch, P. (2001). Lexical typology from a cognitive and linguistic point of view. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *Language typology and language universals: An international handbook* (Vol. 2) (pp. 1143–1175). Berlin: Walter de Gruyter <https://doi.org/10.1515/9783110171549.2.11.1142>
- Koptjevskaja-Tamm, M. (2008). “A lot of grammar with a good portion of lexicon”: Towards a typology of partitive and pseudo-partitive nominal constructions. In J. Helmbrecht, Y. M. Shin, S. Skopeteas, & E. Verhoeven

- (Eds.), *Form and function in language research. Papers in honour of Christian Lehmann* (pp. 329–346). Berlin: Mouton de Gruyter.
- Koptjevskaja-Tamm, M. (2015). Introducing “the linguistics of temperature”. In M. Koptjevskaja-Tamm (Ed.), *The linguistics of temperature* (pp. 1–40). Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.107.01kop>
- Koptjevskaja-Tamm, M. (2016). “The lexical typology of semantic shifts”: An introduction. In P. Juvonen, & M. Koptjevskaja-Tamm (Eds.), *The lexical typology of semantic shifts* (pp. 1–20). Berlin: Mouton de Gruyter.
- Koptjevskaja-Tamm, M., Rakhilina, E., & Vanhove, M. (2015). The semantics of lexical typology. In N. Riemer (Ed.), *The routledge handbook of semantics* (pp. 434–454). London & New York: Routledge.
- Kortmann, B. (1997). *Adverbial subordination: A typology and history of adverbial subordinators based on European languages*. Berlin: Mouton.
- Lakoff, G. (1987). *Women, fire and other dangerous things: What categories reveal about the mind*. Chicago & London: The University of Chicago Press.
- Langacker, R. W. (1988). A view of linguistic semantics. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics* (pp. 49–90). Amsterdam: Benjamins.
- Lehmann, C. (2004). Funktionale Grammatikographie. In W. Premper (Ed.), *Dimensionen und Kontinua. Beiträge zu Hansjakob Seilers Universalienforschung* (pp. 147). 165, Bochum: N. Brockmeyer.
- Levinson, S., Meira, S., & the Language and Cognition Group (2003). ‘Natural concepts’ in the spatial topological domain—Adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language*, 79(3), 485–516. <https://doi.org/10.1353/lan.2003.0174>
- Lichtenberk, F. (1991). Semantic change and heterosemy in grammaticalization. *Language*, 67, 475–509.
- Liljgren, H., & Haider, N. (2015). Facts, feelings and temperature expressions in the Hindukush. In M. Koptjevskaja-Tamm (Ed.), *The linguistics of temperature* (pp. 440–470). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/tsl.107.14lil>
- List, J.-M., Mayer, T., Terhalle, A., & Urban, M. (2014). *CLICS: Database of cross-linguistic colexifications*. Marburg: Forschungszentrum Deutscher Sprachatlas (Version 1.0, retrieved from <http://clics.lingpy.org> accessed on 2017–7-6.
- Luján, E. R. (2010). Semantic maps and word formation: Agents, instruments, and related semantic roles. *Linguistic Discovery*, 8(1), 162–175. <https://doi.org/10.1349/ps1.1537-0852.a.349>
- Luraghi, S. (2001). Syncretism and the classification of semantic roles. *STUF - Language Typology and Universals*, 54(1), 35–51. <https://doi.org/10.1524/stuf.2001.54.1.35>
- Luraghi, S. (2014). Plotting diachronic semantic maps: The role of metaphor. In S. Luraghi, & H. Narrog (Eds.), *Perspectives on semantic roles* (pp. 99–150). Amsterdam: Benjamins. <https://doi.org/10.1075/tsl.106.04lur>
- Majid, A., Boster, J. S., & Bowerman, M. (2008). The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition*, 109(2), 235–250. <https://doi.org/10.1016/j.cognition.2008.08.009>
- Majid, A. M., Bowerman, M., van Staden, M., & Boster, J. S. (2007). The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics*, 18, 133–152.
- Malchukov, A., & Narrog, H. (2009). Case polysemy. In A. Malchukov, & A. Spencer (Eds.), *The Oxford handbook of case* (pp. 518–535). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199206476.013.0035>.
- Malchukov, A., & Ogawa, A. (2011). Towards a typology of impersonal constructions. In A. Siewierska, & A. Malchukov (Eds.), *Impersonal constructions: A cross-linguistic perspective* (pp. 19–56). Amsterdam: Benjamins. <https://doi.org/10.1075/slcs.124.02mal>
- Malchukov, A. L. (2004). Towards a semantic typology of adversative and contrast marking. *Journal of Semantics*, 21(2), 177–198. <https://doi.org/10.1093/jos/21.2.177>
- Malchukov, A. L. (2010). Analyzing semantic maps: A multifactorial approach. *Linguistic Discovery*, 8(1), 176–198. <https://doi.org/10.1349/ps1.1537-0852.a.350>

- Matras, Y. (2004). Typology, dialectology and the structure of complementation in Romani. In B. Kortmann (Ed.), *Dialectology meets typology* (pp. 277–304). Berlin: Mouton de Gruyter.
- Matras, Y. (2009). *Language contact*. Cambridge: Cambridge University Press.
- Mauri, C. (2010). Semantic maps or coding maps? Towards a unified account of the coding degree, coding complexity, and coding distance of coordination relations. *Linguistic Discovery*, 8(1), 210–232. <https://doi.org/10.1349/ps1.1537-0852.a.351>
- Miestamo, M., Bakker, D., & Arppe, A. (2016). Sampling for variety. *Linguistic Typology*, 20(2), 233–296. <https://doi.org/10.1515/lingty-2016-0006>
- Mohammadirad, M., & Rasekh-Mahand, M. (2017). Instrumental semantic map in Western Iranian languages. *STUF - Language Typology and Universals*, 70(4), 579–610. <https://doi.org/10.1515/stuf-2017-0025>
- Narrog, H. (2010). A diachronic dimension in maps of case functions. *Linguistic Discovery*, 8(1), 233–254. <https://doi.org/10.1349/ps1.1537-0852.a.352>
- Narrog, H. (2012). *Modality, subjectivity, and semantic change*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199694372.001.0001>
- Narrog, H., & Ito, S. (2007). Re-constructing semantic maps: The comitative-instrumental area. *STUF - Language Typology and Universals*, 60(4), 273–292. <https://doi.org/10.1524/stuf.2007.60.4.273>
- Narrog, H., & van der Auwera, J. (2011). Grammaticalization and semantic maps. In H. Narrog, & B. Heine (Eds.), *The Oxford handbook of grammaticalization* (pp. 318–327). Oxford: OUP <https://doi.org/10.1093/oxfordhb/9780199586783.013.0025>
- Nikitina, T. (2009). Subcategorization pattern and lexical meaning of motion verbs: A study of the source/goal ambiguity. *Linguistics*, 47(5), 1113–1141. <https://doi.org/10.1515/ling.2009.039>
- Östling, R. (2016). Studying colexification through massively parallel corpora. In P. Juvonen, & M. Koptjevskaja-Tamm (Eds.), *The lexical typology of semantic shifts* (pp. 157–176). Berlin: De Gruyter. <https://doi.org/10.1515/9783110377675-006>
- Perrin, L.-M. (2010). Polysemous qualities and universal networks, invariance and diversity. *Linguistic Discovery*, 8(1), 259–280. <https://doi.org/10.1349/ps1.1537-0852.a.353>
- Perrin, L.-M. (2015). Climate, temperature and polysemous patterns in French and Wolof. In M. Koptjevskaja-Tamm (Ed.), *The linguistics of temperature* (pp. 151–186). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/tsl.107.06per>
- Rakhilina, E. (2015, March). Semantic / lexical typology: Verbs of falling and beyond. Paper presented at the seminar *Verbs, verb phrases and verbal categories*. The Hebrew University of Jerusalem.
- Rakhilina, E., & Reznikova, T. (2014). Doing lexical typology with frames and semantic maps. In *Higher school of economics working papers in linguistics*. Available at <https://ssrn.com/abstract=2564011>
- Rakhilina, E., & Reznikova, T. (2016). A frame-based methodology for lexical typology. In P. Juvonen, & M. Koptjevskaja-Tamm (Eds.), *The lexical typology of semantic shifts* (pp. 95–129). Berlin: De Gruyter.
- Rakhilina, E., Vyrenkova, A., & Plungian, V. (2017). Looking differently: Describing visual direction in Russian and English. Paper presented at the 14<sup>th</sup> International Cognitive Linguistics Conference, Tartu, Estonia.
- Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, 104(4), 1436–1441. <https://doi.org/10.1073/pnas.0610341104>
- Regier, T., Khetarpal, N., & Majid, A. (2013). Inferring semantic maps. *Linguistic Typology*, 17(1), 89–105. <https://doi.org/10.1515/lity-2013-0003>
- Rice, S., & Kabata, K. (2007). Crosslinguistic grammaticalization patterns of the allative. *Linguistic Typology*, 11(3), 451–514. <https://doi.org/10.1515/lingty.2007.031>
- Rijkhoff, J., & Bakker, D. (1998). Language sampling. *Linguistic typology*, 2, 263–314. <https://doi.org/10.1515/lity.1998.2.3.263>
- Ryzhova, D., & Obiedkov, S. (2017). Formal concept lattices as semantic maps. In E. L. Chernyak (Ed.), *Computational linguistics and language science* (pp. 78–87). Aachen CEUR Workshop Proceedings.

- Sandra, D., & Rice, S. (1995). Network analyses of prepositional meaning: Mirroring whose mind—The linguist's or the language user's? *Cognitive Linguistics*, 6(1), 89–130. <https://doi.org/10.1515/cogl.1995.6.1.89>
- Sansò, A. (2010). How conceptual are semantic maps? *Linguistic Discovery*, 8(1), 288–309. <https://doi.org/10.1349/ps1.1537-0852.a.354>
- Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Introduction to multidimensional scaling: Theory, methods, and applications*. New York: Academic Press.
- Schultze-Berndt, E. (2008). What do do verbs do? The semantic diversity of generalised action verbs. In E. Verhoeven, S. Skopeteas, Y.-M. Shin, Y. Nishina & J. Helmbrecht (Eds.), *Studies on Grammaticalization* (pp. 185–208). Berlin: Mouton de Gruyter.
- Siegenthaler, C., & Gunawan, R. (2014). Assessment of network inference methods: How to cope with and underdetermined problem. *PLoS One*, 9(3), e90481. <https://doi.org/10.1371/journal.pone.0090481>
- Siewierska, A., & Papastathi, M. (2011). Third person plurals in the languages of Europe: Typological and methodological issues. *Linguistics*, 43, 575–610.
- Simon-Vandenberghe, A.-M., & Aijmer, K. (2007). *The semantic field of modal certainty: A corpus-based study of English adverbs*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110198928>
- Stassen, L. (1997). *Intransitive predication*. Oxford: Oxford University Press.
- Sweetser, E. (1990). *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Tenser, A. (2008). *The Northeastern dialects of Romani*. (Unpublished doctoral dissertation). UK: University of Manchester.
- Tenser, A. (2016). Semantic map borrowing—Case representation in northeastern Romani dialects. *Journal of Language Contact*, 9(2), 211–245. <https://doi.org/10.1163/19552629-00902001>
- Traugott, E. C. (1985). Conditional markers. In J. Haiman (Ed.), *Iconicity in Syntax* (pp. 289–307). Amsterdam: John Benjamins.
- Traugott, E. C. (2016). Do semantic modal maps have a role in a constructionalization approach to modals? *Constructions and Frames*, 8(1), 98–125. <https://doi.org/10.1075/cf.8.1.07tra>
- van der Auwera, J. (2008). In defense of classical semantic maps. *Theoretical Linguistics*, 34(1), 39–46. <https://doi.org/10.1515/thli.2008.002>
- van der Auwera, J. (2013). Semantic maps, for synchronic and diachronic typology. In A. G. Ramat, C. Mauri, & P. Molinelli (Eds.), *Synchrony and diachrony. A dynamic interface* (pp. 153–176). <https://doi.org/10.1075/slcs.133.07auw>
- van der Auwera, J., Dobrushina, N., & Goussev, V. (2003). A semantic map for imperative-hortatives. In D. Willems, B. Defrancq, T. Coleman, & D. Noël (Eds.), *Contrastive analysis in language. Identifying linguistic units of comparison, Basingstoke* (pp. 44–66). Basingstoke: Palgrave Macmillan. [https://doi.org/10.1057/9780230524637\\_3](https://doi.org/10.1057/9780230524637_3)
- van der Auwera, J., Gast, V., & Vanderbiesen, J. (2012). Human impersonal pronoun uses in English, Dutch and German. *Leuvense Bijdragen*, 98, 27–64.
- van der Auwera, J., Kehayov, P., & Vittrant, A. (2009). Acquisitive modals. In L. Hogeweg, H. de Hoop, & A. Malchukov (Eds.), *Cross-linguistic semantics of tense, aspect and modality* (pp. 271–302). Amsterdam/Philadelphia: John Benjamins.
- van der Auwera, J., & Malchukov, A. (2005). A semantic map for depictive adjectivals. In N. P. Himmelmann, & E. Schultze-Berndt (Eds.), *Secondary predication and adverbial modification. The typology of depictive constructions* (pp. 393–423). Oxford: Oxford University Press.
- van der Auwera, J., & Plungian, V. A. (1998). Modality's semantic map. *Linguistic Typology*, 2(1), 79–124. <https://doi.org/10.1515/lity.1998.2.1.79>
- van der Auwera, J., & Temürçü, C. (2006). Semantic maps. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed., Vol. 11) (pp. 131–134). Oxford: Elsevier.

- van der Auwera, J., & Van Alsenoy, L. (2013a). Vers un cube des oppositions. In J. François, P. Larrivé, D. Legallois, & F. Neveu (Eds.), *La linguistique de la contradiction* (pp. 147–158). Bruxelles: Lang.
- van der Auwera, J., & Van Alsenoy, L. (2013b). Indefiniteness maps: Problems, prospects, and ‘retrospects’. In L. Melazzo (Ed.), *Usare il presente per spiegare il passato. Teorie linguistiche contemporanee e lingue storiche* (pp. 13–26). Il Calamo: Roma.
- Vanhove, M. (2008). Semantic associations between sensory modalities, prehension and mental perceptions: A crosslinguistic perspective. In M. Vanhove (Ed.), *From Polysemy to Semantic Change. Towards a Typology of Lexical Semantic Associations* (pp. 341–370). Amsterdam/Philadelphia: John Benjamins.
- Verkerk, A. (2009). A semantic map of secondary predication. *Linguistics in the Netherlands, 2009*, 115–126. <https://doi.org/10.1075/avt.26.11ver>
- Veselinova, L. (2013). Negative existentials: A cross-linguistic study. *Rivista di Linguistica*, 25(1), 107–145.
- Viberg, Å. (1984). The verbs of perception: A typological study. *Linguistics*, 2, 123–162.
- Wälchli, B. (2010). Similarity semantics and building probabilistic semantic maps from parallel texts. *Linguistic Discovery*, 8(1), 331–371. <https://doi.org/10.1349/ps1.1537-0852.a.356>
- Wälchli, B. (2016). Non-specific, specific and obscured perception verbs in Baltic languages. *Baltic Linguistics*, 7, 53–135.
- Wälchli, B., & Cysouw, M. (2012). Lexical typology through similarity semantics: Toward a semantic map of motion verbs. In M. Koptjevskaja-Tamm & M. Vanhove (Eds.), *New Directions in Lexical Typology Linguistics* (Vol. 50) (pp. 671–710). <https://doi.org/10.1515/ling-2012-0021>
- Youn, H., Sutton, L., Smith, E., Moore, C., Wilkins, J.F., Maddieson, I., Croft, W., & Bhattacharya, T. (2016). On the universal structure of human lexical semantics. *Proceedings of the National Academy of Sciences of the United States of America*, 113(7), 1766–1771. Retrieved from [www.pnas.org/cgi/doi/10.1073/pnas.1520752113](http://www.pnas.org/cgi/doi/10.1073/pnas.1520752113)
- Zwarts, J. (2010a). Semantic map geometry: Two approaches. *Linguistic Discovery*, 8(1), 377–395. <https://doi.org/10.1349/ps1.1537-0852.a.357>
- Zwarts, J. (2010b). Commentary on de Haan—Can semantic maps be built purely bottom-up? *Linguistic Discovery*, 8(1), 123–124. <https://doi.org/10.1349/ps1.1537-0852.a.368>

**Thanasis Georgakopoulos** is Marie Curie's postdoctoral fellow at the University of Liège. He holds a BA in Greek Philology (2004) and an MA (2007) as well as a PhD in Linguistics from the University of Athens (2011). In his dissertation entitled “Diachronic Semantic Analysis of the Ancient Greek Allative Preposition *eis*: A Cognitive Perspective,” he traces the semantic development of *eis* relying on the analytical tools of historical and cognitive linguistics. His current research interests focus on historical semantics, lexical typology, semantic maps, linguistics of space, event perspectivation, and Construction Grammar approach to the semantics of motion verbs. In the past, he was employed as a postdoctoral fellow at the Aristotle University of Thessaloniki, the Free University of Berlin (Excellence Cluster 264 Topoi), and the University of Kassel. As of 2012, he has taught courses on semantics, pragmatics, historical linguistics, cognitive linguistics, and Linguistics and Archeology at the Aristotle University of Thessaloniki, the Humboldt University of Berlin, the University of Kassel, and the University of Münster.

**Stéphane Polis** is a Research Associate at the National Fund for Scientific Research (Belgium) since 2012 and the President of the Department of Ancient Studies at the University of Liège (2016–present). After an MA in classical studies (2002) and an MA in Oriental studies (2004) at the University of Liège, he was awarded a Master of Advanced Studies in general linguistics (2005). He combined the philological and linguistics dimensions in his



PhD entitled “Modality in Late Egyptian” (Liège, 2009). His fields of research are ancient Egyptian linguistics, and Late Egyptian philology and grammar. His work focuses mainly on language variation and language change in ancient Egyptian, with a special interest for the functional domain of modality, grammaticalization processes, and diachronic lexical semantics. He supervises the development of the annotated corpus of ancient Egyptian *Ramses Online* at the University of Liège (with Jean Winand).

**How to cite this article:** Georgakopoulos T, Polis S. The semantic map model: State of the art and future avenues for linguistic research. *Lang Linguist Compass*. 2018;12:e12270. <https://doi.org/10.1111/lnc3.12270>