CHAPTER 15

ON THE STRUCTURE OF TOPONYMS

FRANCESCO-ALESSIO URSINI

Abstract

This paper offers an account of toponyms (e.g. *London*, *Shanghai*, *Roma*, *Kulosaari*), in four languages on a cline of increasing morphological complexity: English, Mandarin, Italian and Finnish. The common morphological processes that underlie these four sets of toponyms are connected to their syntactic distribution and semantic interpretation. An account is thus offered within a minimal version of Type-Logical syntax, coupled with a simple semantic analysis.

Keywords

toponyms, compounds, Type-Logical Syntax, Type-Logical Composition, spatial prepositions.

1. Introduction

Toponyms or place names have received relatively little attention in the theoretical literature (e.g. Anderson 2007, Köhnlein 2015). In English, toponyms have a restricted syntactic distribution, when they occur as complements of spatial prepositions (SPs). Other NPs usually can distribute with articles and quantifiers (e.g. *every*), in complement position. Both types of NPs, qua SPs' complements, denote the landmark object or ground that licenses the spatial relation that an SP denotes (Talmy 2000, 7-14). This fact holds irrespective of the number of located entities or figure(s) that a subject NP can denote (e.g. *the boys*)[1]. These facts are illustrated via (1)-(2). The symbol * represents that the presence of a given lexical item renders a sentence ungrammatical:

[1]Appellative constructions can refer to locations sharing their toponym (e.g. *in every city called "Paris"*: Langendonk 2007, 20-25). We do not discuss them in this paper for mere reasons of space.

(1) *The boys are sitting in front of *∅/the/every car*
(2) *The boys live North of ∅/*the/*every Northampton*

The distributional properties of toponyms seem to have a cross-linguistic import. We illustrate this fact via four languages on a cline of increasing morphological complexity: English, Mandarin, Italian and Finnish. Consider thus (3)-(5):

(3)	*Nánháimen*	*zai*	*∅/*dou*	*Xiǎo xī shān*
	The boys	ZAI	∅/*every	Xiǎo xī shān
	'The boys are in Xiǎo xī shān/*every			Xiǎo xī shān'
(4)	*I ragazzi*	*sono*	*a ∅/*ogni*	*Città di Castello*
	The boys	are	at ∅/*every	Città di Castello
	'The boys are at Città di Castello/*every			Città di Castello'
(5)	*Pojat ovat*	*∅/*jokaisen-GEN*		*Kulosaaren ulkopuolella*
	The boys are	∅/*every-GEN		Kulosaari-GENoutside ADE
	'The boys are outside Kulosaari/*every			Kulosaari'

In (3), the Mandarin toponym *Xiǎo xī shān* distributes with the SP *zai* 'in, at', but not with the universal quantifier *dou* 'all'. In (4)-(5), the same pattern is observed with the Italian toponym *Città di Castello* and Finnish *Kulosaari*. Crucially, morphemes such as *xī* 'west' or *Città* 'city' act as cues for speakers that these NPs are interpreted as toponyms. This is the case for English (Leidner and Lieberman 2011), Mandarin (Huang 2006), Italian (Buscaldi 2010) and Finnish (Arvilia 2014). Thus, the morphological structure of toponyms seems to govern their syntactic distribution and semantic interpretation. However, an account connecting the morphology, syntax and semantics of toponyms is still missing.

The goal of this chapter is to sketch an account with such a unifying perspective. In section 2, we discuss a broader set of toponym data, and the *desiderata* for this account. In sections 3 and 4, we respectively offer the morphological, syntactic and semantic analyses, before concluding.

2. The data

2.1. English toponyms

For English, we focus on the amply documented British toponyms. British toponyms are usually treated as subordinative compounds (Cameron 2003

3-17, Watts 2004, 4-14). Their heterogeneous morphological structures and spelling can be traced to the many substrates of English (e.g. Celt, Latin, and Norse). Both left- and right-headed compounds are attested (e.g. *Aber-deen*, *Edin-burgh*, respectively) with right-headed compounds including a sub-set based on so-called axial nouns (e.g. *North-ampton*, Levinson 1994, Svenonius 2006, 53-54). A third sub-set includes toponyms in which the classifier morpheme can occur in either position, e.g. *kirk* in *Fal-kirk, Kirk-stead*. The head in these compounds is usually labelled a "spatial classifier" morpheme, since it marks the spatial content of a toponym (Aikhenvald 2000, 55-58, Watts 2004, 13-14). Another sub-set includes SPs in their structure, which act as infixed heads (e.g. *Stratford-upon-Avon*, Watts 2004, 15-17). Mono-morphemic or structurally opaque toponyms, usually the result of phonological impoverishment, abound (e.g. *London* from *Londin-ium*).

Examples of these types are offered in (6)-(10):

(6) Left-headed type = {*aber-* 'river' (*Aber-deen*), …}
(7) Right-headed type = {*-burgh/-bury* 'fort' (*Edin-burgh*), *-ampton* 'settlement' (*North-ampton*),…}
(8) Sym-headed type = {*kirk-* 'church' (*Fal-kirk, Kirk-stead*),…}
(9) SP-headed type = {*-upon-* (*Stratford-upon-Avon),…*}
(10) Opaque type = {*London*,…}

Thus, Old English toponyms present an important diachronic picture. The production rules that formed them are now unproductive, but play a role for (word) stress assignment, and interpretation as toponyms. Productive morphological patterns, instead, regulate the emergence of new toponyms, usually considered as semantically transparent phrasal compounds ("compound nouns" in Watts 2004, 19, cf. Scalise and Bisetto 2009). Spatial classifiers are standardly used to create new toponyms (e.g. *street* in *Osbourne street*, Watts 2004, 18-22). Non-spatial classifiers can also be part of a toponym structure, one example being the comparative adjective *greater* in *greater London.* Classifiers can also occur in partitive constructions, which can include a toponym as the head of a partitive and a classifier as its specifier, e.g. *the city of London.* Thus, toponyms in English display the existence of a wealth of non-productive and productive formation patterns[2].

[2] Our analysis of toponyms can be extended to *demonyms* (population names: *Londoner*), *oronyms* and *hydronyms* (e.g. *the Alps*, *the North Sea*, respectively), which usually involve appellative constructions (Anderson 2007, Langendonck 2007). We leave these data aside, also for reasons of space.

2.2. Mandarin toponyms

Mandarin is considered an isolant language (Li and Thompson 1981). However, Mandarin nouns display complex compounding patterns, with a stable mapping between syllabic and morphemic structure (Ceccagno and Basciano 2008). Mandarin toponyms are often poly-morphemic and with a regular structure, possibly related to the lack of non-Mandarin sub-strata (Wu 2006, Zhou *et al*. 2011). For instance, the toponym 小西山 can be transliterated as *Xiǎo xī shān* 'small west hill'. Descriptive morphemes can include adjectives (e.g. *Ān* in *Ān dìng* 'stable city'), and classifiers can include sub-types (e.g. *dà* 'big, large' and *zhōu* 'state'). Many toponyms also include axial nouns, known as *fāngwèicí* in the literature ('localisers': Zhou *et al*. 2011, 42-43). Two examples are *Běijīng* 'North Capital' and *Xiǎo xī shān* 'small west hill'. Mandarin toponyms seem to correspond to subordinative, right-headed and possibly opaque compounds.

Non-exhaustive lists of Mandarin descriptive and classifier morphemes, with matching toponyms, are in (11)-(12), respectively:

(11) Desc. Morph. = {*xiǎo* 'small' (*Xiǎo xī shān*), *ān* 'stable' (*Ān dìng*), *bei* 'North' (*Běijīng*),…}
(12) Class. M. = {*-jing* 'capital' (*Běijīng*), *Zhōu* 'state' (*Ghunag-zhou*),…}

The growth of Chinese cities has brought a need for thousands of street and subway stops toponyms, usually involving poly-morphemic structures (e.g. *Běi tǔ chéng* 'North earth city', Zhou *et al*. 2011). New and old toponyms seem to emerge via the same derivational processes, and have the properties of phrasal compounds (Huang 2006). From a morpho-phonological perspective, tone patterns and stress suggest that toponyms form a single phonological unit. Their semantics might be opaque (e.g. *Běijīng* as 'North capital'), but this is usually not the case for new toponyms (Huang 2006, Wu 2006). Productive and non-productive patterns of toponym formation are thus nearly identical, in Mandarin.

2.3. Italian toponyms

Italian is usually considered a fusive language (e.g. Lorenzi, Salvi and Cardinaletti 2001). Italian toponyms also emerge from the interaction of several substrates (e.g. Latin, Germanic, Greek, Marcato 2011). Suffixation seems a standard process. For instance, toponyms ending in *-ato*, *-ito* or *-eto* can be traced to Latin suffixes that marked a generic location via a

salient feature (e.g. *Pin-eto* 'pine wood'). Pre-nominal spatial classifiers such as *gualdo* or *Fara* (e.g. *Gualdo Tadino*, *Fara Sabina*) are also amply attested.

A sub-set of Italian toponyms includes the preposition *di* 'of' (e.g. *Bassano del Grappa*), with truncated forms being common in spoken language (e.g. *Bassano*). This preposition usually fuses with the definite article to form a complex preposition or *preposizione articolata* (e.g. *Luco de-i Marsi*, Rizzi 1988). The role of *di* in toponyms extends to toponyms for small locations. Furthermore, a small sub-set of Italian toponyms displays an opaque structure including SPs, e.g. *Introdacqua* (from *inter duas acquis* 'between two rivers', Marcato 2011 §4).

A non-exhaustive list of toponym types is presented in (13)-(16):

(13) Suffix type ={*-eto/-ito* 'wood' (*Pineto*, *Coppito*), *sud* 'south',...}
(14) Prefix type ={*gualdo* 'castle' (*Gualdo Tadino*), *piazza* 'square',...}
(15) "Di" type ={*Castel del Monte*,…}
(16) Conflated type ={*Introdaqua*,…}

The suffix type in (13) includes axial nouns (e.g. *nord*), which can have a limited distribution as spatial classifiers (e.g. *Roma nord* 'North Rome'). These morphological patterns double as productive rules: they are employed for names of suburbs and other smaller locations (De Felice 1987). Spatial classifiers are usually found as heads in left-headed compounds, and forms with or without *di* are attested (e.g. respectively *piazza Mandela*, *Borgo degli Elfi*, Marcato 2011). Overall, productive and non-productive patterns are also tightly related, in Italian toponyms.

2.4. Finnish toponyms

Finnish is usually considered an agglutinative language (Karlsson 1999). Finnish toponyms can involve spatial classifiers and spatial case markers as suffixes (Ainiala, Saarelma and Sjölon 2012, 21-27). Spatial classifiers are attested in most toponyms, one example being *-la* 'place' (e.g. *Kumpu-la* 'hill place'). Although this suffix bears a resemblance to the modern adessive case, which ends in *-lla* (e.g. *tallo-lla* 'on top/near the house'), its status as a remnant of a previous locative case is controversial (Aikio 2010). Other examples include *-loki* 'river', *-ranta* 'shore', and other classifiers found in hydronyms and city toponyms.

Suffixation patterns involving case markers are rarer, but nevertheless attested. A small class of toponyms includes infixed genitive case markers, with examples being *Joe-n-suu*, *Helsi-n-ki*, and *Lapee-n-ranta*. A not uncontroversial claim is that these toponyms present the fusion of a genitive case marker as an infix (i.e. *-n-*), with two morphemes, one of them being a spatial classifier (Aikio 2010). A small set of toponyms seems to include the now unproductive essive case as a pure locative case, e.g. *Nuppulin-na* and *Hami-na* (Aikio 2010, 170-172). Other case markers are rarer, so we leave them aside.

Toponyms for streets and minor locations tend to follow the same morphological rules found for other toponyms, suggesting that these processes are productive (e.g. *Kulo-saari*, lit. 'Cod island'). Furthermore, infixed genitive markers seem to partake in productive processes that create street toponyms, but as infixes. Street names often include the combination of a descriptive name, genitive marker and the classifier *katu* 'street' (e.g. *Fleming-in-katu* 'Fleming's street', Ainiala 2012). A list of toponyms involving these affixes is in (17)-(18):

(17) Suffix types = {*-la* 'place' (*Ranti-la*), *-na* 'essive' (*Hami-na*)…}
(18) Genitive types = {*Helsinki, Fleming-in-katu*,…}

The regularity of Finnish toponyms is also mirrored in their semantics, as they invariably refer to specific locations, viz. (5). Thus, both spatial classifiers and spatial case markers play a role in diachronic and synchronic processes of toponym formation, and may still be distinguishable in opaque toponyms (e.g. *-la* in *Rantila*).

Overall, our compact but precise discussion suggests that toponyms across these four languages involve both productive and non-productive compounding and affixation rules. These rules combine descriptive morphemes/words (nouns and adjectives) and spatial classifiers, case markers and suffixes. Regardless of the morphological processes at stake, toponyms share the same distributional properties and interpretation, viz. (1)-(5). Thus, our goal is to capture these rules under a unified account.

3. The analysis: the morphology of toponyms

3.1. TLS and the architecture of grammar

The framework we adopt is a minimal version of *Type Logical Syntax* (TLS: e.g. Carpenter 1992, Moortgat 2010). We choose this formal

framework, since it emphasizes a transparent syntax-semantics interface, and a derivational system that captures how morpho-syntactic structures are derived. Although TLS is neutral to questions of grammar architecture, we follow a perspective in which the boundaries between morphology and syntax are blurred (cf. Halle and Marantz 1993, Keenan and Faltz 2004).

We start by defining our basic building blocks. In standard TLS, morpho-syntactic categories are mapped or assigned onto *types*, which are represented as being either "complete" or "incomplete" information units. Complete types represent lexical items that can stand as independent constituents (e.g. *np* for noun phrases as *the girl*). Incomplete types are assigned to items that must combine with other elements to form a complete type (e.g. *s/np* for the verb *runs*). Two lexical items can be combined or merged if their types match: the sentence *the girl runs* can be formed, because the *input* type of *runs* match the type of *the girl*.

For our analysis of toponyms, we concentrate on the treatment of morphological *features* that toponyms *qua* NPs carry (cf. Carpenter 1992). We thus assign the general type *p* to descriptive morphemes, and the type *p'* to morphemes carrying spatial features. Before we pursue this aspect fully, however, we introduce the connectives / and • to represent the *right division* and the *product* operations (Jäger 2005, Moortgat 2010). We define right division as a *binary*, *associative* operation; product, as a *non-commutative* operation: *x•y* involves the ordered pair of *x* and *y*. Our rules are listed below:

1. *Given a Lexicon* $\boldsymbol{L}$, $\boldsymbol{p}$ *is a morphological type* (Lexical type)
2. *If* $\boldsymbol{x}$ *is a type and* $\boldsymbol{y}$ *is a type, then* $\boldsymbol{x/y}$ *is a type* (Type f.: division)
3. *If* $\boldsymbol{x}$ *is a type and* $\boldsymbol{y}$ *is a type, then* $\boldsymbol{x \bullet y}$ *is a type* (Type f.: product)
4. *If* $\boldsymbol{x/y}$ *is a type and* $\boldsymbol{y}$ *is a type, then* $\boldsymbol{(x/y) \bullet y \vdash x;\ y \bullet (x/y) \vdash x}$ (F.A.)
5. *Nothing else is a type* (Closure Rule)

Rule 1 introduces *p* as the general type, from which other types are defined. Rule 2 defines the formation of functional complex types; rule 3, that of product complex types. Rule 4 is known as "forward application" ("F.A.": Moortgat 2011, §2.1): matching types (e.g. $\boldsymbol{x}$ and x) are "cancelled out", but if they do not match (e.g. we merge $\boldsymbol{x}$ and $\boldsymbol{y}$), a derivation is said to *diverge* or *crash*. Forward application ("F.A.") operates in a distributed manner. The product (•) of two matching types proves the existence of phase with a third type, a result we represent via ⊢. Rule 5 says that no other rules are needed. We do not use other connectives as they not crucial, for our account (e.g. left division \: Morryll 2011).

Finally, we capture the cyclic nature of our derivations by defining a simple pre-order as the pair of an interval set *I*, and an addition operation +, i.e. <*I*,+>. We implement two operations, lexical selection (LS) and merge introduction (MI) to explicitly mark the introduction of a new element in a derivation and the merging of two elements, respectively.

3.2 The analysis: the morpho-syntax of toponyms

Our goal in this section is to capture the morphological structures attested in each language via our TLS system. Therefore, we assume that our derivations capture non-productive and productive toponym formation rules alike. This also holds for the types we assign to morphemes. Even if a morpheme is not a recurring morpheme in a certain language (e.g. *Aber-*, *-deen* in Modern English), our type assignment simply captures how they can contribute to the structure of the toponym *Aberdeen*.

We start by offering an account of the English data. The two types of toponyms that involve right- or left-headedness can be derived by merging a 1-place classifier head with a descriptive morpheme, whether it be an axial noun (e.g. *North* in *Northampton*), or some other descriptive morpheme. Descriptive morphemes are thus assigned the type *p* of phrases or arguments; spatial classifiers, the functional type *p'/p* of 1-place heads. The merge of these two morphemes, irrespective of their linear position, is a phrase of type *p'*, viz. (19)-(21):

(19) *t.* [$_{p'/p}$ Aber] (LS)
t+1. [$_{p}$ deen] (LS)
t+2. [$_{p'/p}$ Aber]•[$_{p}$ deen]⊢ [$_{p'}$ Aberdeen] (MI)

(20) *t.* [$_{p}$ North] (LS)
t+1. [$_{p'/p}$ hampton] (LS)
t+2. [$_{p}$ North]•[$_{p'/p}$ hampton]⊢[$_{p'}$ Northampton] (MI)

(21) a. *t+2.* [$_{p}$ Fal]•[$_{p'/p}$ kirk]⊢[$_{p'}$ Falkirk] (MI)
b. *t+2.* [$_{p'/p}$ Kirk]•[$_{p}$ stead]⊢[$_{p'}$ Kirkstead] (MI)

As (19)-(20) show, *Aberdeen* and *Northampton* include a spatial classifier on either side (i.e. *hampton* 'settlement' and *deen* 'river') of the descriptive morpheme. However, either version of forward application proves that a toponym, an NP carrying the spatial features *p'*, is formed. The compressed derivations for *Falkirk* and *Kirkstead* in (21), instead, suggest that in some cases both versions of forward application *were* at

work in English, when these toponyms emerged. Via TLS, we can capture how both combinations of morphemes are merged, if attested.

Toponyms including SPs as an infixed head require a slightly more complex analysis. We assign type *p'/p'/p'* to SPs such as *upon* in *Stratford-upon-Avon*, and type *p'* its toponym arguments (here, *Avon* and *Stratford*). We thus follow treatments such as Hale and Keyser (2002, 41-54), which assume SPs take other spatial phrases as arguments ("P-within-P" hypothesis). These toponyms present recursive toponym structures, in which an SP takes two toponyms (*Avon* and *Stratford*, type *p'*) as its arguments, and forms a toponym phrase, carrying spatial features and meaning (also type *p'*). This is shown in (22):

(22) *t.* [$_{p'}$ Stratford] (LS)
t+1. [$_{p'/p'/p'}$ upon] (LS)
t+2. [$_{p'}$ Stratford]•[$_{p'/p'/p'}$ upon]⊢[$_{p'/p'}$ Stratford-upon] (MI)
t+3. [$_{p'}$ Avon] (LS)
t+4. [$_{p'/p'}$ Stratford-upon]•[$_{p'}$ Avon]⊢[$_{p'}$ Stratford-upon-Avon] (MI)

For novel toponyms (e.g. *Osbourne street*), we assume that free classifiers such as *street* also act as 1-place heads, as shown in (23):

(23) *t+2.* [$_{p'}$ Osbourne]•[$_{p'/p'}$ street]⊢[$_{p'}$ Osbourne street] (MI)

Overall, the derivations in (19)-(22) show that toponyms can emerge via different derivational processes, which nevertheless involve the same basic categories: descriptive morphemes, spatial classifiers and (possibly) SPs.

We can now turn to the Mandarin data. Given their homogeneous structure, they find a simple explanation in our account, viz. (24)-(25):

(24) *t.* [$_{p'}$ Běi] (LS)
t+1. [$_{p'/p'}$ jīng] (LS)
t+2. [$_{p'}$ Běi]•[$_{p'/p'}$ jīng]⊢[$_{p'}$ Běijīng] (MI)

(25) *t.* [$_{p'}$ Xiǎo] (LS)
t+1. [$_{p'/p'}$ xī] (LS)
t+2. [$_{p'}$ Xiǎo]•[$_{p'/p'}$ xī]⊢[$_{p'}$ Xiǎo xī] (MI)
t+3. [$_{p'/p'}$ shān] (LS)
t+4. [$_{p'}$ Xiǎo xī]•[$_{p'/p'}$ zhān]⊢[$_{p'}$ Xiǎo xī shān] (MI)

As (24)-(25) show, the merge of a descriptive morpheme (e.g. the axial noun *běi* 'north') with a spatial classifier (e.g. *jīng* 'capital') derives a

toponym of type *p'*. Toponyms involving sequences of descriptive morphemes can involve adjectives (e.g. *xī* 'small'), as modifiers of type *p/p*, forming phrases that can merge with spatial classifiers (e.g. *Xiǎo xī shān* 'west small hill'). This analysis can be extended to similar data in the other languages (e.g. *Greater London*), and more generally it can account forms of affix or case stacking, via our rules (cf. Ursini 2014, 2015a).

The picture that emerges for the Italian data, then, follows a similar tack. Consider the compressed derivations in (26)-(29):

(26) *t+2.* [$_{p}$ Copp-]•[$_{p'/p}$ –ito]⊢[$_{p'}$ Coppito] (MI)
(27) *t+2.* [$_{p'/p}$ Gualdo]•[$_{p}$ Tadino]⊢[$_{p'}$ Gualdo Tadino] (MI)
(28) *t+2.* [$_{p'/p}$ Piazza]•[$_{p}$ Mandela]⊢[$_{p'}$ Piazza Mandela] (MI)
(29) *t+4.* [$_{p'/p'}$ Castel del]•[$_{p'}$ Monte]⊢[$_{p'}$ Castel del Monte] (MI)

The derivations in (26)-(28) show that, whether a suffix (*-ito* in *Coppito*) or a spatial classifier (*Gualdo*, *Piazza*) merges with a descriptive morpheme in either position, a toponym of type *p'* is derived. For partitive toponyms, such as *Castel del Monte* in (29), *di* can be treated as an SP that merges with two toponyms as its arguments (*Castel*, *Monte*).[3] The presence or absence of the definite article type seems to hinge on subtle distributional factors, hence we leave a more thorough discussion aside.

We conclude with Finnish toponyms, which require a simple discussion. Consider the derivations in (30)-(31):

(30) *t+2.* [$_{p}$ Kulo]•[$_{p'/p}$ saari]⊢[$_{p'}$ Kulosaari] (MI)
(31) *t.* [$_{p}$ Fleming] (LS)
t+1. [$_{p'/p'/p}$ -in-] (LS)
t+2. [$_{p}$ Fleming]•[$_{p'/p'/p}$ in]⊢[$_{p'/p}$ Fleming-in] (MI)
t+3. [$_{p'}$ katu] (LS)
t+4. [$_{p'/p}$ Fleming-in]•[$_{p'}$ katu]⊢[$_{p'}$ Fleming-in-katu] (MI)

Since toponyms are formed via the suffixation of spatial classifiers or case markers, both categories can be assigned the type *p'/p* of 1-place heads. Either head can then merge with a descriptive morpheme of type *p*, as shown in (30), which includes the spatial classifier *-saari* 'lake' is merged with the descriptive morpheme *Kulo*. We thus blur the distinction between these two categories, a move justified on their distribution and semantics

[3] Since the definite article can be assigned type *p'/p'* of 1-place heads (e.g. Ritter 1993, Szabolcsi 2010), the "cut rule" operation can form a complex head (i.e. we have *(p'/p'/p')•(p'/p')⊢ p'/p'/p'*: Moortgat 2010, §2.1).

(cf. Ursini 2014). For toponyms including genitive case markers as 2-place heads (e.g. *-in-*), our supplementary assumption is that their type is *p'/p'/p*. These markers map a classifier and a descriptive morpheme onto a recursive toponym type. Observe that classifiers such as *katu* display a dual nature, since they can also be (nominal) arguments of this head. For this reason, we assign them type *p'*, as shown in (30). However, spatial classifiers in Italian, English and Mandarin also display these properties (e.g. *la via*, *the street*, *Jiēdào*). These systematic valence patterns can be captured via the so-called residual rule: a morpheme of type *p'* can be raised to type *p'/p'* (Moortgat 2011 §2.1, Ursini 2015a, b). In other words, 0-place heads (i.e. arguments) such as nouns for location types can become spatial classifiers (1-place heads), as they carry spatial features that allow the merge of these heads with another spatial phrase.

Our analysis can now also shed light on why their distribution with quantifiers is blocked for semantic reasons. Here we offer one English example in (32), in which the quantifier *every* is assigned the type *p/p* of 1-place heads (cf. again Ritter 1993, Szabolcsi 2010). Nevertheless, the type of analysis can be safely extended to the other three languages, viz. (3)-(5), given the structural equivalence of the patterns involves:

(32) *t+k.* [$_{p/p}$ every]•[$_{p'}$ London]⊢* (Derivation crashes)

The merge of quantifiers with toponyms involves non-matching input types (i.e. *p'* and *p*); the derivations crashes. Furthermore, our analysis captures that toponyms must merge with SPs. If SPs are of type *p'/p'/p'*, then toponyms are perfect complement phrases for SPs, since their type is *p'*. Other NPs such as *the car* in (1) must be implicitly or explicitly (e.g. Finnish case markers) assigned to type *p'*, to merge with SPs (cf. Zwarts and Winter 2000, Svenonius 2006).

Let us now summarize the key points of our analysis. First, toponyms involve the forward application, or merge of two types of morphemes: descriptive morphemes and spatial 1-place heads. Descriptive morphemes can include axial nouns, nominal adjectives (e.g. *xiǎo* 'small') and nominal roots describing located entities (e.g. *Copp-* in *Coppito* 'poplars' place'). Spatial heads can include classifiers, case markers and suffixes. Note that, irrespective of the orthographic conventions in each language, the morphological processes that derive toponyms are shown to be the same. Second, the merge order of these morphemes is not crucial, but their type is. When quantifiers are merged with toponyms, a mismatch in types renders a phrase, and hence a sentence, ungrammatical.

4. The analysis: the semantics of toponyms

4.1. The apparatus: types for toponyms

The analysis we sketch for the semantics of toponyms is based on type theory and λ-calculus (e.g. Szabolcsi 2010, 3-30), but hopefully the core claims should be easily accessible to all readers. As we seen in sections 1-2, descriptive morphemes denote properties or labels identifying a place (e.g. *xiǎo* 'small'), and spatial classifiers the spatial interpretation of a toponym (e.g. *city*). This spatial interpretation is the central type of sense assigned to toponyms, in SP contexts. Crucially, TLS is based on a transparent mapping between syntax and semantics: morpho-syntactic types determine semantic ones (Jäger 2005, 170-178, Moortgat 2010, §4). Thus, lexical items of type *np* are assigned semantic type *e*, for referents, while type *s/np* items are assigned semantic type $e>t$ of properties.

Since we reason with morphological features, we also implement a slightly more refined theory for semantic types. We use a fragment of *Type Logic Composition* (TLC, Asher 2011), which implements a richer ontology including also locations, events and individuals, and is theoretically close to TLS. Given our analysis of toponyms as denoting locations, we assume that the morphosyntactic type *p* is mapped onto the semantic type *e* (individuals/referents), and type *p'* to semantic type *l* (locations). The use of these types is explained in the next section.

4.2. The analysis: the semantics of toponyms

From the semantic side, the tight relations between symmetric derivational processes are easily captured. *Function Application* holds whether the argument of a function appears to the left or right of a function (Szabolcsi 2010, 14-15). Consequently, spatial classifiers, case markers and affixes are all interpreted as 1-place functions. Since several classifiers have an opaque sense, from a diachronic perspective (e.g. *Aber-* in *Aberdeen*), we use the function $\lambda x.\boldsymbol{class'}(x)_{e\rightarrow l}$ to translate their senses. Spatial classifiers indeed classify the sense of a toponym as belonging to the spatial type, even if its exact interpretation may be opaque to speakers. Classifiers also trigger the specific, definite readings observed for toponyms: one referent is mapped onto its corresponding, specific location (cf. Anderson 2007, 70-75). Consider (33):

(33) *t.* [[North]]$\vDash n_e$ (Int)
t+1. [[hampton]]$\vDash \lambda x.\boldsymbol{class'}(x)_{e\rightarrow l}$ (Int)
t+2. ([[North]])×[[ampton]]$\vDash (n_e)\lambda x.\boldsymbol{class'}(x)_{e>l}=\boldsymbol{class'}(n)_l$ (FA)

We represent the interpreted lexical items via the "Interpretation" function (i.e. [[.]]), and function application as the composition × of function and argument, which yields a λ-conversion (i.e. the relation $\vDash$, cf. Jäger 2005). Thus, the sense of *Northampton* corresponds to a referent that is classified as a location, although the precise sub-type is not accessible.

This analysis can be applied to all the opaque toponyms including a spatial classifier, viz. (19)-(21) for English, (24)-(25) for Mandarin, (26)-(27) for Italian, (30) for Finnish. For toponyms including SPs, e.g. *upon* in English (22), *di* in Italian (29), genitive case markers in Finnish (31), our analysis only requires that these relations are assigned slightly different types. SPs, given our rules of interpretation, denote spatial relations as relations between locations, of type $l\rightarrow(l\rightarrow l)$ (cf. also Zwarts and Winter 2000). Instead, *di* and Finnish genitive case markers denote (underspecified) relations between a description and the location it is associated to (i.e. type $e\rightarrow(l>l)$). Consider (34)-(35):

(34) *t.* [[Stratford]]$\vDash st_l$ (Int)

t+1. [[upon]]$\vDash \lambda x.\lambda y.\boldsymbol{upon'}(x,y)_{l\rightarrow(l\rightarrow l)}$ (Int)

t+2. ([[Stratford]]))×[[upon]]$\vDash(st_l)\lambda x.\lambda y.\boldsymbol{upon'}(x,y)_{l\rightarrow(l\rightarrow l)}$
$=\lambda y.\boldsymbol{upon'}(st,y)_{l\rightarrow l}$ (FA)

t+3. [[Avon]]$\vDash av_l$ (Int)

t+4. [[Stratford-upon]]×([[Avon]])$\vDash\lambda y.\boldsymbol{upon'}(st,y)_{l\rightarrow l}(av_l)$
$=\boldsymbol{upon'}(st,av)_l$ (FA)

(35) *t+k.* [[Fleming]]×[[in]] $\vDash(f_e)\lambda x.\lambda y.R(x,y)_{e\rightarrow(l\rightarrow l}$
$=\lambda y.R(f,y)_{l\rightarrow l}$ (FA)

t+n. [[Fleming-in]]×[[katu]]$\vDash\lambda y.R(f,y)_{l\rightarrow l}(k_l)=R(f,k)_l$ (FA)

The derivations in (34)-(35) show that both types of toponyms (here, *Stratford-upon-Avon* and *Fleming-in-katu*) denote locations, although they do so by combining different types of *relata*. Insofar as the corresponding relation receives the type *l* of locations, the precise spatial relation holding between the two referents can remain underspecified (here, the relation *R*). The same reasoning can apply to Italian toponyms such as *Castel del Monte*. Thus, toponyms can include morphemes that have non-spatial senses. Nevertheless, their morphological structure is interpreted as denoting a location named via some of its possible salient features.

For productive toponyms, the transparent structures they display suggest that their semantic is more easily accessible (e.g. Buscaldi 2010, Zhou *et al*. 2011). Speakers can generally interpret these toponyms as place names for streets, squares and so on. This is reflected in the interpretation we offer for the Mandarin and Italian examples in (36)-(37):

(36) $t{+}4$.[[Xiǎo xī]]×[[zhān]] $\vDash$ $(\textbf{\textit{small'}}(w))_e\lambda x.\textbf{\textit{hill'}}(x)_{e\rightarrow l}$ (FA)
$=\textbf{\textit{hill'}}(\textbf{\textit{small'}}(w))_l$

(37) t. [[Piazza]]×[[Mandela]]$\vDash\lambda x.\textbf{\textit{square'}}(x)_{e\rightarrow l}(m)_e$ (FA)
$=\textbf{\textit{square'}}(m)_l$

The derivation in (36) is based on the assumption that *xī* 'small' acts as a modifier of *xiǎo* 'west', thus acting as function of type $e\rightarrow e$ (cf. Asher 2011, ch. 3). This west, small entity is then classified as a hill via *zhān* 'hill', thus being assigned the type l of locations. Our system, then, indirectly predicts that the presence of several, stacked descriptive morphemes is semantically unproblematic. It also predicts that the linear position of a classifier is irrelevant semantics-wise, as the interpretation of the Italian *Piazza Mandela* in (37) shows.

For quantifiers, then, we assume a simplified semantics with respect to standard approaches (e.g. generalized quantifier theory), but one that brings them closer to their distribution as 1-place heads (cf. Szabolcsi 2010). For instance, *every* denotes a function that takes a property, and applies it to a quantified domain: we have $\lambda P.\forall xP(x)_{(e\rightarrow t)\rightarrow t}$, with t the type of truth-values. Consider (38), the interpretation of (32):

(38) $t{+}k$ [[every]]×([[London]])$\vDash\lambda P.\forall xP(x)_{(e\rightarrow t)\rightarrow t}\times(l)_l$=#(Der. Crash.)

Thus, (38) shows that quantifiers require a semantic type that corresponds to the sense of a common noun, which in turn denotes a property of the quantified entities. Toponyms, in virtue of their interpretation as specific locations (of type l), cannot be composed with quantifiers, which denote sets of individuals (of type $(e\rightarrow t)\rightarrow t$). The wrong semantic type blocks function application, and renders an ungrammatical sentence also uninterpretable, as a consequence.

Overall, our analysis captures how toponyms are interpreted as specific locations, entities of type l, even if their senses may have become opaque. This analysis is tightly related to the morphosyntactic analysis: toponyms receive this interpretation in virtue of the morphological features that they carry. The analysis thus indirectly predicts that toponyms merge with SPs as bare NPs, but cannot merge with quantifiers. We thus have a sketch of how the different linguistics aspects of toponyms are brought together in one analysis. Since we have reached our goal, we can conclude.

5. Conclusions

In this chapter we have sketched a theory of toponyms in four languages: English, Mandarin, Italian and Finnish. Even if these languages can form a cline of increasing morphological complexity, toponyms seem to share similar morphological structures, which guide their syntactic and semantic properties. Our analysis captures these facts via a simple TLS treatment, and a matching TLC semantics. However, it does not cover other aspects of this very complex topic, and must gloss over comparisons with related theories on toponyms (e.g. Langendonk 2007, Köhnlein 2015), as well as theories of onomasiology (e.g. Štekauer 2005). We leave such complex topics for future research.

References

Aikhenvald, Alexandra Y. *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press, 2006.

Aikio, Aante. "The study of the Saami substrate toponyms in Finland." *Ononamstica Uralica* 4 (2010): 159-198.

Ainiala, Terhi, Minna Saarelma, and Paula Sjölon. *Names in Focus: An Introduction to Finnish Onomastics*. Suomalaeisen Kirjallisuuden Seura, 2012.

Ainiala, Terhi. "Place names and identities: the uses of names of Helsinki." *Oslo Studies in Language*, 4 (2) (2012): 7-15.

Anderson, John M. *The Grammar of Names.* Oxford: Oxford University Press, 2007.

Asher, Nicholas. *Lexical Meaning in Context: A Web of Words*. Oxford: Oxford University Press, 2011.

Buscaldi, Davide. "Toponym ambiguity in geographical information retrieval." In *Proceedings of the 32nd international ACM SIGIR Conference on Research and Development in Information Retrieval*, edited by James Allan and Javed Aslam, 7-12. Boston, MA: ACM Digital Library, 2010.

Cameron, Ken. *A Dictionary of British Place Names*. Manchester: Manchester University Press, 2003.

Carpenter, Ken. *The Logic of Typed Feature Structures*. Cambridge: Cambridge University Press, 1992.

Ceccagno, Alessandra and Bianca Basciano. "Classification of Chinese compounds," In *Online Proceedings of the Sixth Mediterranean*

Morphology Meeting, edited by Angela Ralli, Geert Booij, Sergio Scalise and Athanasios Karasimos, 71-83. Patras: University of Patras, 2008.

De Felice, Emidio. Onomastica. In *Linguistica Storica*, edited by Roberto Lazzeroni, 147-179. Roma: La Nuova Italia Scientifica, 1987.

Hale, Ken and Samuel J. Keyser. *Prolegomena to a Theory of Argument Structure*. Cambridge, MA: The MIT Press, 2002.

Halle, Morris and Alec Marantz. "Distributed morphology and the pieces of inflection." In *The View from Building* 20, edited by Ken Hale and Samuel J. Keyser, 111-176. Cambridge, MA: The MIT Press, 1993.

Huang, Degen. "Automatic recognition of Chinese toponyms." *Computer Engineering* 32 (2006): 220-222.

Jäger, Gerhard. *Anaphora and Type Logical Grammar*. Dordrecht: Springer, 2005.

Karlsson, Fred. *A Grammar of Finnish*. London: Routledge, 1999.

Keenan, Edward and Edward Stabler. *Bare Grammar*. Stanford, CA: CSLI Press, 2004.

Köhnlein, Björn. "The morphological structure of complex place names: the case of Dutch." *Journal of Comparative Germanic Linguistics*, 18 (1) (2015): 183-212.

Langendonck, Willy van. *Theory and Typology of Proper Names*. Berlin and New York: Mouton de Gruyter, 2007.

Leidner, Jonathan and Michael D. Lieberman**. "**Detecting geographical references in the form of place names and associated spatial natural language." *SIGSPATIAL Special*, 3 (2) (2011): 5-11.

Levinson, Stephen C. "Vision, shape and linguistic description: Tzeltal body-part terminology and object description." *Linguistics*, 32 (4/5) (1994): 791-856.

Li, Charles N. and Sandra Thompson. *Mandarin Chinese: A Functional Reference Grammar*. Los Angeles, CA: University of California Press, 1981.

Lorenzi, Lorenzo, Giampaolo Salvi and Anna Cardinaletti, (eds.). *Grande Grammatica di Consultazione*. Bologna: Il Mulino, 2001.

Marcato, Carla. "Toponimi." In *Grande Enciclopedia dell'Italiano* (2nd edn.), edited by Raffaele Simone. Bologna: Treccani, 2011. Accessed December 1st, 2015, http://www.treccani.it/enciclopedia/toponimi_(Enciclopedia-dell'Italiano)/

Moortgat, Michael. J. "Type-Logical grammar." In *The Stanford Encyclopedia of* Philosophy, edited by Edward N. Zalta, nn.pp. Stanford, CA: The metaphysics research lab, 2010. Accessed

December 1st, 2015, http://plato.stanford.edu/entries/typelogical-grammar/.

Ritter, Erica. "Where's gender?" *Linguistic Inquiry*, 24 (4) (1993): 795-803.

Rizzi, Luigi. "Il sintagma preposizionale." In *Grande Grammatica Italiana di Consultazione*, 3, edited by Lorenzo Renzi, 497-530. Bologna: Il Mulino Editore, 1988.

Scalise, Sergio and Antonella Bisetto. "The classification of compounds." In *The Oxford Handbook of Compounding*, edited by Rochelle Lieber and Pavol Štekauer, 34-53. Oxford: Oxford University Press, 2009.

Štekauer, Pavel. *Meaning Predictability in Word Formation*. Amsterdam and Philadelphia: John Benjamins, 2005.

Svenonius, Peter. "The emergence of axial parts." *Nordlyd*, 33 (1) (2006): 49-77.

Szabolczi, Anna. *Quantification*. New York, NY: Basil Blackwell, 2010.

Talmy, Leonard. *Towards a cognitive Linguistics*. Cambridge, MA: The MIT press, 2000.

Ursini, Francesco-Alessio. "The morphology of spatial P: Is a unitary perspective possible? " In *Proceedings of Les Decembrettes 8*, edited by Sandra Augendre *et al*., 371-398. Toulouse: Mirail.

Ursini, Francesco-Alessio. "On the syntax and semantics of Italian spatial Ps. " *Acta Linguistica Hungarica*, 63 (1) (2015a): 3-51.

Ursini, Francesco-Alessio. "Another look at preposition stranding. " In *A Festschrift for Nils-Lennart Johansson*, edited by Britt Erman, Philip Shaw, Gunnel Melchers, & Peter Sudnkvist, 319-343. Stockholm: Stockholm University Press, 2015b.

Watts, Victor. *The Cambridge Dictionary of English Place-Names: Based on the Collections of the English Place-Name Society*. Cambridge: Cambridge University Press, 2004.

Wu, Zhirong. "Discussion on toponym wording." *Cartography*, 4 (1) (2006): 42-43.

Zhou, Dongmei, Maohua Qian, Min Hua, Dan Liu, and Xuri Tang. "Structural analysis and computation of Chinese toponyms." *International Journal of Knowledge and Language Processing*, 2 (3) (2011): 36-47.

Zwarts, Joost and Yoad Winter. (2000). "Vector space semantics: a model-theoretic analysis of locative prepositions." *Journal of Logic, Language and Information*, 9 (2) (2000): 169-211.