

RESEARCH ARTICLE

Malmö, Simrishamn, and Norrland: asymmetry in the categorization of place

Marlene Johansson Falck  and Kristina Persson 

Department of Language Studies, Umeå University, 901 87 Umeå, Sweden

Corresponding author: Marlene Johansson Falck; Email: marlene.johansson.falck@umu.se

(Received 16 January 2024; revised 2 November 2024; accepted 6 November 2024)

Abstract

Categorization plays a crucial role in organizing experiences, allowing us to make sense of the world. This process is reflected in the labels speakers use for geographical areas. This study investigates the categorization of geographical areas reflected in phrases including nouns for the three Swedish regions of *Norrland*, *Svealand*, or *Götaland*, and the conjunction *och* ('and'). Using data from the Swedish Korp corpus (Borin et al. 2012), we examine how these regions and areas within them are represented in governmental, news, and social media texts. Results show that *Svealand* and *Götaland* are more commonly used with nouns for regions than *Norrland*. *Norrland* is used with phrases for more specific locations within the other regions (e.g. their towns and provinces) but also considerably larger areas (e.g. countries and continents) more commonly than the other regions, revealing asymmetry in how geographical areas in Sweden are categorized.

Keywords: asymmetry; categorization; coordination; geographical categories; *Götaland*; *Norrland*; *Svealand*

1. Introduction

People commonly sort their experiences into kinds (Lakoff 1987):

Every time we see something as a *kind* of thing, for example, a tree, we are categorizing. Whenever we reason about *kinds* of things – chairs, nations, illnesses, emotions, any kind of thing at all – we are employing categories. Whenever we intentionally perform any *kind* of action, say something as mundane as writing with a pencil, hammering with a hammer, or ironing clothes, we are employing categories. (Lakoff 1987:5–6)

Categorization processes such as these are grounded in our body–world interactions and essential for making sense of the world (Lakoff 1987). By means of them, we organize our experiences into categories (here defined as a number of entities that

are considered equivalent; see Rosch 1978:30), which are designated by the terms that speakers use (Rosch, 1978:30). Categories may be sorted into cognitive taxonomies (i.e. *kind-of*-hierarchies based on *kind-of* relations; chairs and desks are kinds of furniture [Mark et al. 1999:285], and Sweden and Norway are kinds of countries) and partonomies (i.e. *part-of* hierarchies based on *part-of* relations; seconds are parts of minutes, minutes are parts of hours, and hours are parts of days [Mark et al. 1999:285]; Stockholm is part of Svealand, Svealand is part of Sweden, Sweden is part of Europe). However, entities that are part of the same geographical category are not necessarily categorized the same way by the speakers of a language. Consider the following *och Norrland* sentences from the three subcorpora of governmental, news, and social media texts (henceforth abbreviated GT, NT, and SMT¹) from the Swedish Korp corpus (Borin et al. 2012). They all include the noun² *Norrland*, which denotes one of Sweden's three main regions, Norrland, Svealand, and Götaland.

- (1) Från vartdera Malmö, Simrishamn och Norrland
from each Malmö, Simrishamn and Norrland
 fanns det endasten [sic] representant.
was there onlyone [sic] representative
 'There was only one representative from Malmö, Simrishamn, and Norrland.'
 (GT, *och Norrland*)
- (2) Hos SJ är sug-et störst efter biljetter
with SJ is the demand biggest after tickets
 från Stockholm till Skåne och Jämtland,
from Stockholm to Skåne and Jämtland,
 men även till Göteborg, Sundsvall, Falun-Mora och Norrland.
but also to Gothenburg, Sundsvall, Falun-Mora and Norrland.
 'With SJ, the demand for tickets is the biggest [on trains] from Stockholm to Skåne and Jämtland, but also to Gothenburg, Sundsvall, Falun-Mora and Norrland.' (NT, *och Norrland*)
- (3) Man vet att man är stockholmare när Sverige består av
you know that you are Stockholmer when Sweden consists of
 Stockholm, skåne [sic] Göteborg, Värmland, Dalarna och Norrland.
Stockholm, skåne, Göteborg, Värmland, Dalarna and Norrland.
 'You know you are a Stockholmer (i.e. from Stockholm) when [you consider] Sweden to consist of Stockholm, Skåne, Göteborg, Värmland, Dalarna and Norrland.' (SMT, *och Norrland*)

In sentence (1), *Norrland* is coordinated with nouns for towns within Götaland (*Malmö* and *Simrishamn*). The sentence specifies from what towns in Götaland there was only one representative (of something) without providing the same level of detail in the information about the representative from a place in Norrland. All we learn is that he or she is from *Norrland*. Sentence (2) reflects more complex and, indeed, incorrect categorization. Here, *Norrland* is coordinated with nouns for

towns in Svealand (*Stockholm* and *Falun-Mora*), Götaland (*Göteborg*), and Norrland (*Sundsvall*) and with nouns for counties in Götaland (*Skåne*) and Norrland (*Jämtland*). From this sentence, we learn that the demand for train tickets with SJ is the biggest between towns and counties (within or outside of Norrland) and the region of Norrland. It suggests to the speaker that 'Norrland' is something other than that of the Norrland town of *Sundsvall* and the Norrland province of *Jämtland*. In sentence (3), *Norrland* is again coordinated with the nouns for the province of *Skåne* and the town *Göteborg* in Götaland, and with the town *Stockholm* in Svealand, but also with the counties or provinces *Värmland* and *Dalarna* in Svealand. It again shows that *Norrland* may be coordinated with nouns for more specific geographical categories within the other two regions of Sweden. As in sentence 1, places in Norrland are not categorized as instances of the more specific categories they are part of (e.g. as towns and counties) but sorted into the larger category of Norrland. Furthermore, from the sentence, we learn that to the speaker, the actual marks of a person from Stockholm are to consider Sweden to consist of the towns of *Stockholm* and *Göteborg*, the counties or provinces of *Värmland* and *Dalarna*, and then *Norrland*. In that sense, it provides information on the speaker's attitude towards these geographical categories. When considering some geographical areas, categorization at the levels of towns (Stockholm and Göteborg) or counties or provinces (Värmland and Dalarna) is essential. However, regarding *Norrland*, categorization at the region level appears to be sufficient. Specific places within this region are not categorized as such but sorted into the broader category of Norrland.

Sentences (1)–(3) thus suggest a potential asymmetry in how geographical categories are categorized, particularly in how Norrland and the other two main regions of Sweden are categorized. Although *Norrland*, *Svealand*, and *Götaland* refer to the same geographical category (i.e. to regions), they do not seem to be treated equally.

This study investigates the categorization of geographical locations, as reflected in phrases including the nouns *Norrland*, *Svealand*, and *Götaland* and the Swedish conjunction *och* ('and'). The study is based on corpus linguistic analyses of up to 1,000 random instances of sentences including each of the phrases *och Norrland*, *och Svealand*, *och Götaland*, *Norrland och*, *Svealand och*, or *Götaland och* from the GT, NT, and SMT subcorpora of the Swedish Korp corpus (Borin et al. 2012). The purpose of the study was to test the hypothesis that the asymmetries reflected in sentences (1)–(3) are by no means unusual. Analyses of the representations of Norrland in contemporary newspaper texts (Eriksson 2010) have shown that Norrland is discursively constructed as an abstract space with certain specific connotations. However, we do not know how common such usage patterns are and whether Norrland and areas within Norrland differ from Svealand and Götaland and areas within these regions in this regard.

The following research question were asked:

- RQ1. Are *Svealand* and *Götaland* more commonly used with nouns for one of the three main regions compared with Norrland?
- RQ2. Is *Norrland* more often used with nouns for geographical categories within regions than *Svealand* and *Götaland*?

We expected *Norrland* to be coordinated with nouns for regions less commonly than *Svealand* and *Götaland*, but more frequently with phrases for geographical places within regions.

The paper is structured in the following way. A short background to the regions of Sweden and a discussion of the functions of the nouns *Norrland*, *Svealand*, and *Götaland* is provided in Section 2, and a theoretical background to research on cognitive categorization processes in Section 3. Section 4 deals with the materials and method used and Section 5 the results. We conclude with a discussion of the results, limitations of the study, and prospects for future research in Section 6.

2. The regions of Norrland, Svealand, and Götaland

Sweden is the fifth largest country in Europe, and with its elongated form (1,500 km) similar in length and size to California (<https://sweden.se/life/society/key-facts-about-sweden>). It has twenty-five provinces (Sw. *landskap*) historically grouped into three regions (Sw. *landsdelar*): Norrland, Svealand, and Götaland (Teleman et al. 1999). The entities of *landskap* and *landsdel* do not fill any administrative functions today. Instead, *riksområde* is used for statistical purposes in relation to the territorial units in the European Union, NUTS (Nomenclature des Unités Territoriales Statistiques) (Förordning (EU) 2019/2088 av den 27 november 2019 om hållbarhetsrelaterade upplysningar som ska lämnas inom den finansiella tjänstesektorn 2019), and *län* for administrative purposes. Still, references to the historical regions are commonly made in news and weather reports.

The climate in the north of Sweden is sub-arctic with long, cold winters and short summers, while in the south, the climate is semi-boreal with a damper climate and milder winters with few days of snow. However, Sweden has a milder winter climate than other areas on the same latitude (<https://www.smhi.se/kunskapsbanken/klimat/sveriges-klimat/sveriges-klimat-1.6867>).

Sweden has a long coastline on the east and southwest coasts, while the Scandes (Sw. *Skanderna*) mountain range delimits the country to Norway in the northwestern part. As much as 69% of Sweden is covered with forests (<https://sweden.se/life/society/key-facts-about-sweden>). Despite its size, Sweden is one of Europe's most sparsely populated countries (with 21 inhabitants per square kilometer). Still, most Swedes live in urban areas³ (Sw. *tätorter*). According to SCB (<https://www.scb.se/hitta-statistik/sverige-i-siffror/miljo/tatorter-i-sverige/>), 72% of inhabitants in the least populated province of Norrland live in such areas and 83% of inhabitants in the most populated province. The equivalent numbers for Svealand are 83% in the least populated province, and 97% in the most populated one, and for Götaland 64% and 91%, respectively.

The urban areas are primarily located along the Swedish coasts or around the lakes of Mälaren, Hjälmaren, Vättern, or Vänern. Norrland and Svealand are particularly unevenly populated, with substantially smaller numbers of inhabitants inland. The three biggest municipalities in Sweden are its capital Stockholm (with 1.6 million inhabitants), Göteborg (with 608,000 inhabitants), and Malmö (with around 325,000 inhabitants).

Norrland is the largest and northernmost of the three regions but has the smallest population. It covers about 58% of Sweden (261,292 square kilometers), and about 12% of the Swedish population (about 1.17 million inhabitants) live there. Still, six of Sweden's thirty largest urban areas are in Norrland, two of which have more than 100,000 inhabitants (<https://www.scb.se/hitta-statistik/sverige-i-siffror/miljo/tatorter-i-sverige/>). They are generally located along the coast, while areas in the western part of Norrland are more sparsely populated.

Svealand is the smallest of the regions and is located between Norrland and Götaland. It covers about 20% of Sweden (91,098 square kilometers), and about 40% of Swedes (about 4 million people) live there. Ten of the thirty most populated municipalities are in Svealand. Eight of them have more than 100,000 inhabitants (<https://www.scb.se/hitta-statistik/sverige-i-siffror/miljo/tatorter-i-sverige/>). Several of them are located around the lake Mälaren. Historically, Svealand was Sweden's administrative center (Asker 2009), which is also reflected in contemporary Sweden.

Götaland is the southernmost and most densely populated of the regions. It covers about 22% of Sweden (97,841 square kilometers), and about 48% of Sweden's inhabitants (about 4.78 million people) live there. Nearly half of the thirty most populated areas in Sweden are in Götaland. Nine of them have more than 100,000 inhabitants (<https://www.scb.se/hitta-statistik/sverige-i-siffror/miljo/tatorter-i-sverige/>).

The nouns *Norrland*, *Svealand*, and *Götaland* are all proper names, which prototypically have unique referents (Teleman et al. 1999:117). This feature distinguishes proper names from appellatives, and therefore proper names are sometimes treated as a separate part of speech. In addition to grammatical differences, the main differences between proper names and nouns are semantic. Proper names, like nouns, have referents, but they are labels for unique referents rather than referents with functional traits or qualities in common. Different places may have the same place name even if they do not share any features besides those related to topography. Such meanings have often faded but are obvious in nature names with suffixes such as *-vik* ('bay') and *-berg* ('mountain'). Ainala & Östman (2017:3) imply that etymology is largely irrelevant when names are used in context. A much-debated question reflecting the distinctive feature between proper names and appellatives is whether proper names have meaning or sense in addition to reference (for an overview see Helleland 2012:99).

Because place names are not associated with semantic meaning like common nouns, but primarily have a referential function, it is not always clear what referents the *Norrland*, *Svealand*, and *Götaland* instances have. The reference of a place, or speakers' conceptions of this reference, may be very wide, as in *Sweden*, or very narrow, as in cultural names like *Pekkantalo* ('the homestead of Pekka'). *Norrland*, *Svealand*, and *Götaland* denote large regions, and speakers might not always have the whole regions in mind when using them but may use the terms metonymically for more specific areas that make up parts of the regions (e.g. for towns and provinces within them). Such broad categorization may be imprecise (as reflected in sentence 1), incorrect (as in 2), or sweeping, as in (3), and involve asymmetry in speakers' categorization of place. Furthermore, it is not always clear, or agreed upon, where the borders of some geographical areas really are. For example, where do mountains such as Kebnekaise and Mount Everest begin and end, and where were

the borders of Rome in 410 AD? Moreover, geographical names may be used in reference to several overlapping areas and several different concepts. Previous research has shown that *Norrland* is not only used in reference to a Swedish region but about a combination of geographies, histories, climates, borders, and ideas (Eriksson 2010:26). First, in territorial terms it is equal to the sum of Sweden's five northernmost counties (Eriksson 2010:28). Second, it is used for a somewhat larger area north of the climatological border, *Limes Norlandicus* (Sw. *biologiska norrlandsgränsen*) (Rune 1995). Third, cultural boundaries between Norrland and the rest of Sweden can be drawn by mapping patterns in material cultural heritage, such as settlements and artifacts, and immaterial ones, such as language varieties. The concept of *kulturgräns* ('cultural border') by the ethnologist Sigurd Erixon (Edlund & Frängsmyr 1995), has been developed in, for instance, the project *Kulturgräns norr* (Edlund & Greggas 2000). Fourth, the concept of *Norrland* and speakers' attitudes towards this region are discursively constructed. Eriksson (2010) has analyzed popular representations of the Swedish north in contemporary news articles about Norrland retrieved from the database *Presstexts*. The study is based on articles from the daily newspaper *Dagens Nyheter* from the years 2000–2004 retrieved in two searches, one with the keywords *Jämtland*, *Gävleborg*, *Norrbottnen*, *Västerbottnen*, and *Västernorrland* and one with the keyword *Norrland*. Using critical discourse analysis tools, she shows that the texts presuppose that Norrland is a region in need of support, with migration from the area and a surplus of men, without critically revising the claims made. Similarly, its counties are constructed as sparsely populated areas without acknowledging variations within them (Eriksson 2010:52). Her analysis (Eriksson 2010:57) of whether the articles with the keyword *Norrland* further specify the specific areas involved (e.g. counties, municipalities, or cities) shows that only seventeen out of thirty-eight articles provide this information. When the areas are specified, the information typically comes at the end of texts referencing Norrland. She suggests that in this way, the areas are represented as familiar but abstract spaces with well-known connotations. Furthermore, specified areas primarily involve the northernmost counties (Norrbottnen and Västerbottnen) and use specific aspects of them to represent the whole county. Symptomatically, the province of Dalarna (a mountainous province with traditional rural areas that is a part of Svealand) is more often labeled *Norrland* than the Norrland coastal province of Gävleborg (Eriksson 2010:57). Eriksson (2010:57) thus argues that journalistic practices of framing the area reinforce stereotypes. Through them, certain aspects of the region become universalized, and Norrland becomes a place that lacks variation. Contemporary representations of Norrland are related to the broader and much discussed dichotomy between center and periphery (Eriksson 2010:25), and historical issues related to societal questions in the comparatively poor north that were much debated when the northern part of Sweden became industrialized during the late nineteenth and early twentieth centuries (Edlund & Frängsmyr 1995:293). Because the natural resources in Norrland were essential to industrialization and hence played an important part in the development of modern Sweden, Norrland was then framed as 'Framtidslandet' ('The Land of the Future', Sörlin 2023).

To conclude, place names are not only expressions denoting specific places. Besides their referential function, they have social and cultural functions (Ainiala &

Östman 2017:3). They vary in style according to situation, reflecting speakers' conceptions of and attitudes towards such places. Speakers also have feelings attached to certain places. Helleland (2012:109) reflects upon place identity, arguing that:

names are abstractions of the places they refer to, substituting physical features with a wide range of expressions . . . place names constitute links between name users and the named objects, as well as with the past.

In this sense, place names also have a symbolic function.

3. Cognitive categorization processes

Early work on categorization processes includes Rosch's (1973) work on natural categories, and Medin and Schaffer's (1978) work on exemplar-based reasoning. Rosch demonstrated that category judgments are based on comparisons between new instances and perceptually salient natural prototypes of each category, and Medin and Shaffer (1978) showed that comparisons with exemplars (i.e. remembered examples of a category) and the contexts of these memories are also involved when people form categories.

It is well established that linguistic labels impact categorization (Althaus & Plunkett 2016:770). For instance, one study by Lupyan et al. (2007) has shown that linguistic labels can facilitate category learning. In the study, participants were first taught to categorize 'alien' as 'those to be approached' or 'those to be avoided' (Lupyan et al. 2007:1007), and then given accuracy feedback that was either accompanied by nonsense labels of the categories presented or by no linguistic label at all. The study showed that even though all participants had been given the same accuracy feedback and the labels were essentially redundant, those given labels for the categories they were shown were better at learning the categories than those who had not. Yet other studies have shown that linguistic labels can impede the identification of visual categories (e.g. Robinson & Sloutsky 2007) and that labels can induce a focus on commonalities between the entities that are part of a given category. One example of the latter finding is Althaus and Plunkett's (2016) study of the eye movement of 12-month-olds while these infants were familiarized with the categories of leaves and shells. The study showed that labels influence how categories are processed. Infants presented with labels for the objects they were shown noticed their commonalities. In contrast, those not shown such labels focused on the variability between the objects displayed.

The level of categorization that is most basic to human experience (henceforth *basic-level category*) includes objects and organisms such as dogs, cats, cars, buses, chairs, and sofas (Rosch 1978:30), which may be included in larger, more inclusive and hence more abstract *superordinate categories* (e.g. dogs and cats may be categorized as kinds of animals, cars and buses as kinds of vehicles, and chairs and sofas as kinds of furniture), or subdivided into smaller, less inclusive, more specific *subordinate categories* (e.g. dogs may be subdivided into the categories of Alsatians and Poodles, cars into the categories of Volvo and Porsche, and chairs into

armchairs and office chairs). However, categorization processes are not neutral. Rosch (1978) argued that unexpected precision in the way objects are categorized (e.g. referring to one's *pencils* as one's *Mongol number 3 pencils*) can have a snobbish or satirical effect, and Cruse (1977) argued that lower degrees of specificity de-emphasize features that are omitted while higher degrees of specificity emphasize the features that are added by a more specific category.

The parts of geographical paronyms are all basic-level categories. Still, the labels speakers use (e.g. *Sweden*, *Svealand*, and *Stockholm*) about the same geographical area (e.g. Stockholm) may differ in their specificity. In one case, the categories employed may be fine-grained and specific (e.g. refer to towns, or parts of towns), and in another, broad and imprecise (e.g. refer to regions, countries, or continents).

Based on the results of empirical investigations such as those above, we take the nouns that speakers use for geographical places to reflect speakers' categorization of these places and mediate how they categorize these places to others. Next, we describe the methods and materials employed in our analysis of categorization reflected in speakers' uses of the phrases *och Norrland*, *och Svealand*, *och Götaland* and *Norrland och*, *Svealand och*, and *Götaland och*.

4. Materials and method

The OCH REGION subsets were excerpted on 18 February 2023 from the GT, NT, and SMT subcorpora of the Swedish Korp corpus (Borin et al. 2012). The subcorpora were chosen because they include diverse types of texts written by different speakers of Swedish and are intended for different audiences. GT is the most official subcorpus with texts of Swedish agencies and authorities by representatives of the public, NT the most public one with texts in standard, written Swedish by skilled writers (journalists) for large readerships, and SMT the most informal one with texts written by people expressing their views on topics online. Given these differences, GT provides information on categorization by speakers of Swedish in official contexts, NT on categorization by speakers in public contexts, and SMT on categorization by speakers typically representing themselves.

At the time of excerpting the data, GT comprised twenty-five subcorpora and around 128.8 million sentences, NT fifty-nine subcorpora and 54.9 million sentences, and SMT sixty-nine subcorpora and 843.5 million sentences. The REGION OCH subsets were excerpted from the same subcorpora on 10 June 2024. GT then comprised thirty-five subcorpora and 175.8 million sentences. NT still comprised fifty-nine corpora, but 55.4 million sentences, and SMT sixty-nine 69 corpora and 862 million sentences.

4.1 Data collection and preparation of data

We excerpted 1,000 random instances of each phrase with more than 1,000 instances from each subcorpus, and all instances of phrases with less than 1,000 instances in the subcorpus. These processes resulted in nine subsets of *och Norrland*, *och Svealand*, and *och Götaland* (henceforth OCH REGION subsets) comprising a total of 5,818 instances, and nine subsets of *Norrland och*, *Svealand och*, and *Götaland*

och (REGION OCH subsets) comprising 9,000 instances. The frequency differences reflect the fact that there were less than 1,000 instances in the *och Svealand* and *och Götaland* subsets.

Next, we removed duplicates, near-duplicates (i.e. nearly identical instances from the same source text, e.g. SVT NYHETER 2014), and instances that did not make sense graphically or content-wise (e.g. '10.8 14.0 16.3 13.4 20.3 ... och Norrland' [GT *och Norrland*]) from the subsets. Furthermore, instances with syntactic forms that do not involve coordination of phrases within a clause (i.e. with any of the nouns *Norrland*, *Svealand*, or *Götaland*) were excluded. They are instances where *och* ('and') coordinates two or more main clauses. One example is (4), in which two main clauses are coordinated with the main clause of *Norrland har enorma tillgångar* ('Norrland has enormous resources').

- (4) Vi ska se till att hela landet ska leva – det är
 we will see to that whole country shall live – this is
 mycket viktigt – och Norrland har enorma tillgångar.
 very important – and Norrland has enormous resources.

'We will see to it that the whole country will live – this is very important – and Norrland has enormous resources.' (MT *och Norrland*)

Taken together, 14% of the OCH REGION subsets and 20% of the REGION OCH subsets were excluded in the preparation of the data described above.

4.2 Analysis of data

Next, we identified the geographical categories coordinated with *Norrland*, *Svealand*, or *Götaland* in the remaining instances by analyzing the heads of the phrases coordinated with each of the *och Norrland*, *och Svealand*, *och Götaland*, *Norrland och*, *Svealand och*, and *Götaland och* instances. The analyses were done individually by each researcher followed by meetings in which the analyses were compared and discrepancies discussed.

Some of the *och Norrland* (22%), *och Svealand* (3%), and *och Götaland* (11%) instances were coordinated with phrases for several geographical locations. Because they were occasionally very complex, but typically involved coordination with phrases for the same type of geographical location,⁴ they were categorized in accordance with the first phrase.

4.2.1 Main categories

The following main categories were identified.

CONTINENT: For instance, *Afrika* ('Africa'), *Sydamerika* ('South America'), and *Europa* ('Europe'). In (5), *Asia* and *Norrland* are coordinated.

- (5) Finnair har gjort bedömningen att det finns behov
Finnair has done the assessment that there is need
av enklare och snabbare förbindelser mellan Asien och Norrland.
of simpler and faster connections between Asia and Norrland.
'Finnair assessed that there is a need for simpler and faster connections between
Asia and Norrland.' (NT och Norrland)

COUNTRY: For example, phrases such as *Sverige* ('Sweden') and *riket* ('the country') or *landet* ('the country') implying that Sweden is the country, but also phrases for other countries (e.g. *Kanada* 'Canada' or *Turkiet* 'Turkey') coordinated with *Norrland*, *Svealand*, or *Götaland*.

- (6) Doreen har bland annat bott i Nigeria, Zambia och Norrland.
Doreen has among other lived in Nigeria, Zambia and Norrland.
'Doreen has lived in Nigeria, Zambia and Norrland among other places.'
(NT och Norrland)

PROVINCE: Phrases that designate the contemporary administrative units of *län* ('county') or the historical but still vital units of *landskap* ('province'). They were sorted into one category partly because some names (e.g. *Gotland*) are used for both *län* and *landskap*, and partly because the relationship between *län* and *landskap* is geographically intricate. In some cases, a *län* is a smaller part of one or several *landskap*, for example *Kronobergs län* (the southwestern part of *landskapet Småland*) and *Stockholms län* (which is part of both *landskapet Uppland* and *landskapet Södermanland* geographically). In yet others, *län* consists of more than one *landskap*: *Norrbottnens län* and *Västerbottnens län* consist of the *landskap Norrbotten* and *Västerbotten* respectively, and the *landskap of Lappland* which is divided between the two.

- (7) Som jag ser det genomsyras hela Skåne, Småland, Västergötland
as I see it permeates whole Skåne, Småland, Västergötland
och Norrland av denna bonniga attityd.
and Norrland of this hillbilly attitude.
'As I see it the whole of Skåne, Småland, Västergötland and Norrland is
permeated with this hillbilly attitude.' (SM och Norrland)

REGION: *Norrland*, *Svealand*, and *Götaland*.

- (8) Det avgörs på onsdag och kommun-en tävlar då
it is decided on Wednesday and the municipality competes then
mot de båda vinnarna i Götaland och Svealand.
against them both winners in Götaland and Svealand.
'It will be decided on Wednesday when the municipality competes against the
winners in Götaland and Svealand respectively.' (NT och Svealand)

AREAS: Phrases for somewhat vaguely delimited areas such as *Mälardalen* (around the lake *Mälaren* in *Svealand*), *Bergslagen* (the elongated geographical area

in Svealand traditionally associated with mining and metallurgic industry), and *Norrlandskusten* ('the coastal area of Norrland').

PLACES: Cities, towns, and municipalities (e.g. *Stockholm*, *Örebro*, and *Vetlanda*).

ISLANDS AND LAKES (e.g. *Öland* and *Vänern*).

OTHER GEOGRAPHICAL ENTITIES: Phrases for geographical categories that do not fall into any of the above geographical categories. They include phrases for ISLANDS, LAKES, OR SEAS outside of Sweden (e.g. *Phuket*, *Medelhavet* 'the Mediterranean'), DIRECTIONS (e.g. *söderut* 'south', *uppåt* 'up'), and AREAS such as *skogsbygd* ('forest area') and *glesbygd* ('sparsely populated area'), which do not have an exact location but may be in, or spread over, one or more of the regions. This category also includes historical names of geographical areas in or outside contemporary Sweden.

4.2.2 Subcategories

Some of the *Norrland*, *Svealand*, and *Götaland* instances are coordinated with single nouns (e.g. *Gotland* and *Vetlanda*) and others with attributes modifying and hence delimiting the area denoted by the noun. These more delimited areas make up more SPECIFIC geographical categories and were thus labeled as such. In (9), the description of *Asia* is narrowed down to *South Asia* and therefore categorized as CONTINENT-SPECIFIC.

- (9) Detta är nog något Apple själva har märkt och
this is probably something Apple themselves have noticed and
 tänkt på när iPad har tagit över värld-en och man ser
thought on when iPad has taken over the world and you see
 dem i afrikanska byar och småkiosker i Sydosien [sic] och
them in African villages and small kiosks in South Asia and
 Norrland och ja, dator-n får nog överlag stå tillbaka.
Norrland and well, the computer may generally stand back.

'This is something that Apple have noticed and thought about when iPad has taken over the world and you see them in African villages and small kiosks in South Asia and Norrland and well, the computer may generally have to stand back.'
 (SMT och Norrland)

In the same way, the adjective *södra* ('south') specifies which parts of Sweden are denoted by *södra Sverige* ('south of Sweden') and the noun phrase *östra delarna av* ('the eastern parts of') which parts of the regions of Svealand and Götaland are denoted by the phrase *östra delarna av Svealand och Götaland* ('the eastern parts of Svealand and Götaland'). Instances including the phrases *södra Sverige* and *östra delarna av Svealand och Götaland* are thus categorized as COUNTRY-SPECIFIC and REGION-SPECIFIC, respectively. The categories of AREA-SPECIFIC, PROVINCE-SPECIFIC, and PLACE-SPECIFIC are based on the same logic.

Table 1 summarizes the identified geographical categories sorted in a continuum from the largest to the smallest categories, with region in the middle.

Table 1. Continuum of geographical categories combined with Norrland, Svealand, or Götaland

Area	Geographical category
Larger than regions	CONTINENT CONTINENT-SPECIFIC COUNTRY COUNTRY-SPECIFIC
Region	REGION
Areas within regions	REGION-SPECIFIC PROVINCE PROVINCE-SPECIFIC AREAS PLACE PLACE-SPECIFIC ISLANDS AND LAKES
Other geographical categories	ISLANDS, LAKES, and SEAS outside Sweden DIRECTIONS

4.2.3 Instances excluded from the study

In the analysis, two categories were identified that were not relevant for the focus and hence excluded from the study. First, because our focus was on coordination with *Norrland*, *Svealand*, or *Götaland*, and not with phrases for more delimited areas of these regions, we removed *och Norrland*, *och Svealand*, and *och Götaland* instances that do not denote the whole region. One example is (10) in which the attributive phrase *östra* ('eastern') modifies and hence delimits not only the area denoted by the coordinated noun (*Götaland*) but also, in an elliptic way, the *Svealand* instance retrieved from the subcorpus.

- (10) Varmast blir det i östra Götaland och Svealand
the warmest becomes it in eastern Götaland and Svealand
 där det kan bli 15 grader på måndag.
where it can be 15 degrees on Monday.
 'It will be the warmest in eastern Götaland and Svealand on Monday.' (NT,
och Svealand)

For the present study, we decided that attributive phrases denoting adjacent areas, such as *västra Svealand och Götaland*, modify both coordinated nouns, whereas those denoting areas further apart, e.g. *södra Götaland and Norrland*, do not. Instances such as these were found in the REGION OCH subset. They were the most common in the *Svealand och* instances (16%), the second most common in the *Götaland och* instances (8%), and the least common in the *Norrland och* instances (2%), suggesting that *Svealand och* and *Götaland och* instances are more prone to being preceded by an attributive phrase delimiting the area denoted by the noun than *Norrland och*.

Second, because our focus was on GEOGRAPHICAL categories, phrases for NON-GEOGRAPHICAL categories (e.g. *snö* 'snow', *vinter* 'winter', *sommar* 'summer', *fiske*

'fishing', and *bönder* 'farmers') denoting weather, seasons, leisure activities, and people were also excluded. These instances were primarily found in the *Norrland* instances (207 instances), which may be compared with thirty-five *Svealand* instances and fourteen *Götaland* instances, suggesting that there are differences between *Norrland*, *Svealand*, and *Götaland* also regarding this type of data.

5. Results

Figures 1–3 show the proportions of phrases for geographical categories coordinated with the remaining *och Norrland*, *och Svealand*, and *och Götaland* instances in the subsets excerpted from GT, MT, and SMT; Figures 4–6 show those in the *Norrland och*, *Svealand och*, and *Götaland och* instances. The categories include nouns for the REGIONS of *Norrland*, *Svealand*, and *Götaland*, phrases for areas within the regions (i.e. REGION-SPECIFIC areas, PROVINCES, PLACES, PLACE-SPECIFIC areas, AREAS WITHIN REGIONS, and LAKES AND ISLANDS in Sweden), phrases for larger geographical areas (i.e. CONTINENTS, COUNTRIES, and CONTINENT-SPECIFIC and COUNTRY-SPECIFIC areas), and phrases for OTHER AREAS (e.g. GEOGRAPHICALLY VAGUE AREAS in Sweden, SEAS, ISLANDS, and AREAS outside Sweden).

The proportions of phrases for geographical categories coordinated with the OCH REGION subsets (Figures 1–3) are very similar. In all three subsets, *Svealand* (93% in GT, 90% in NT and SMT) and *Götaland* (87% in GT, 89% in NT, 68% in SMT) are coordinated with phrases for the three main regions considerably more often than *Norrland* is (14% in GT, 28% in NT, 15% in SMT) (RQ1). Moreover, *Norrland* is coordinated with phrases for areas within regions considerably more often (67% in GT, 56% in NT, 66% in SMT) than *Svealand* (7% in GT, 9% in GT and SMT) and *Götaland* (6% in GT, 12% in NT, 7% in SMT) (RQ2). Furthermore, the results show that *Norrland* (14% in GT, 13% in NT and SMT) is coordinated with phrases for areas that are larger than regions more frequently compared to *Svealand* (0% in GT, 1% in NT and SMT) and *Götaland* (1% in GT and NT, 2% in SMT).

The proportions of phrases for geographical categories in the REGION OCH subsets (Figures 4–6) similarly show that *Svealand* (37% in GT and NT, 68% in SMT) and *Götaland* (68% in GT, 51% in NT, 76% in SMT) are coordinated with phrases for the regions more often than *Norrland* (2% in GT, 15% in NT, 6% in SMT) (RQ1). Moreover, *Norrland* is coordinated with phrases for areas within regions more often than *Svealand* in SMT (66% vs. 23%) and with *Götaland* in all three subsets (51% vs. 26% in GT, 53% vs. 45% in NT, 66% vs. 20% in SMT) (RQ2). However, the proportion of *Svealand och* instances in GT and NT (57% in both subsets) coordinated with phrases for areas within regions is higher than both the *Norrland och* instances (51% in GT, 53% in NT) and *Götaland och* instances (25% in GT, 45% in NT). Although a very small proportion of *Norrland och* instances (2% in GT, 15% in NT, 6% in SMT) are coordinated with phrases for REGIONS and primarily with phrases for areas within regions, coordination with *Norrland och* involves a large proportion of instances coordinated with phrases for other categories apart from these two (47% in GT, 32% in NT, 28% in SMT). Moreover, coordination with *Svealand och* in GT and NT is substantially different from coordination with *och Svealand* and *Svealand och* in SMT.

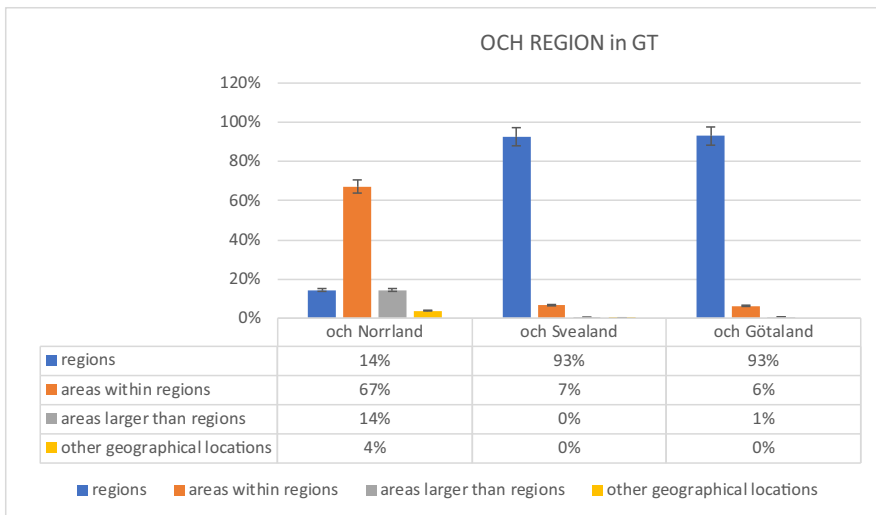


Figure 1. OCH REGION in governmental texts.

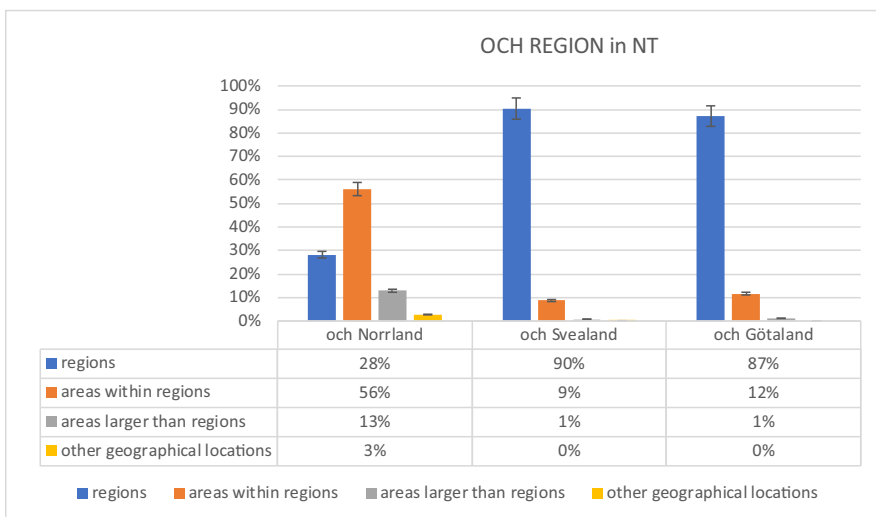


Figure 2. OCH REGION in news texts.

Taken together, the proportions of phrases in Figures 1–6 confirm the hypothesis that *Svealand* and *Götaland* are more often coordinated with nouns for the three main regions than *Norrland*, and those in four out of six subsets confirm the hypothesis that *Norrland* is coordinated with phrases for geographical places within the regions of Sweden more frequently compared with *Svealand* and *Götaland*. The results highlight asymmetries in how *Norrland*, *Svealand*, and *Götaland* and areas within these regions are categorized, but also show that *Svealand* is commonly followed by phrases for areas within regions in GT and NT.

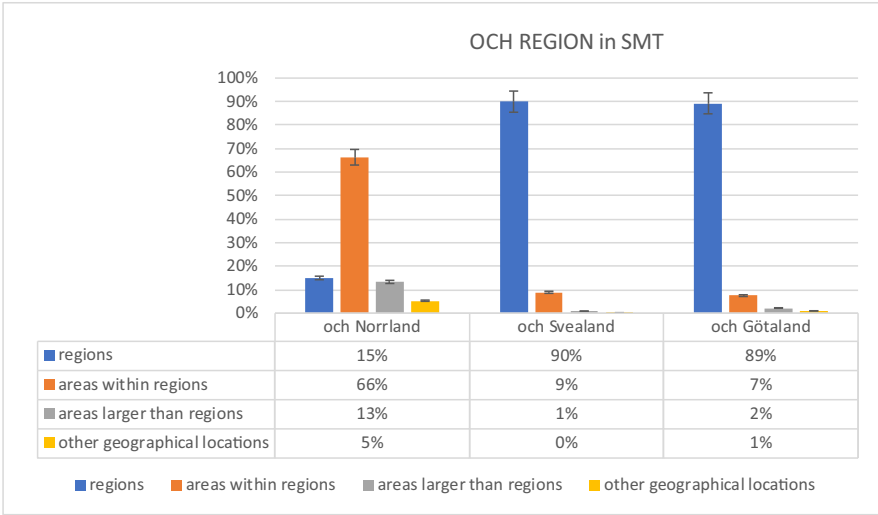


Figure 3. OCH REGION in social media texts.

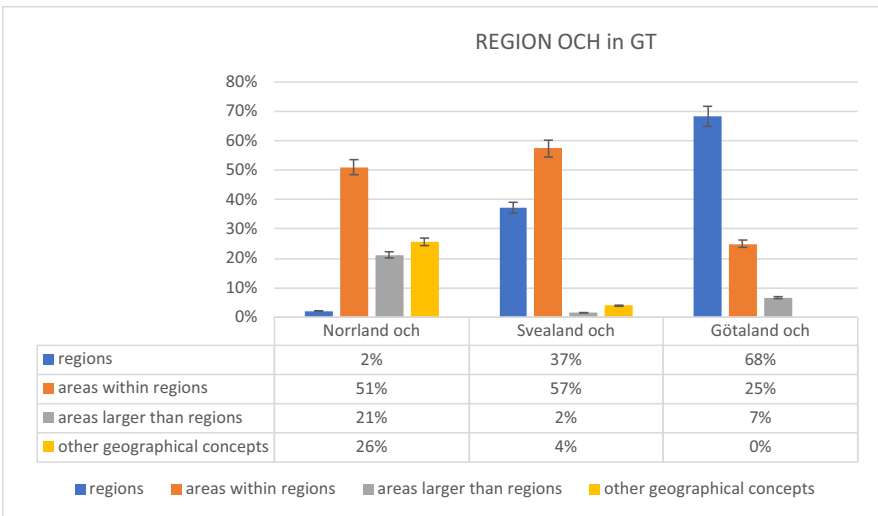


Figure 4. REGION OCH in governmental texts.

Sections 5.1–5.4 discuss the usage patterns of coordination with phrases for REGIONS, AREAS WITHIN REGIONS, AREAS LARGER THAN REGIONS, and OTHER AREAS in the OCH REGION and REGION OCH subsets.

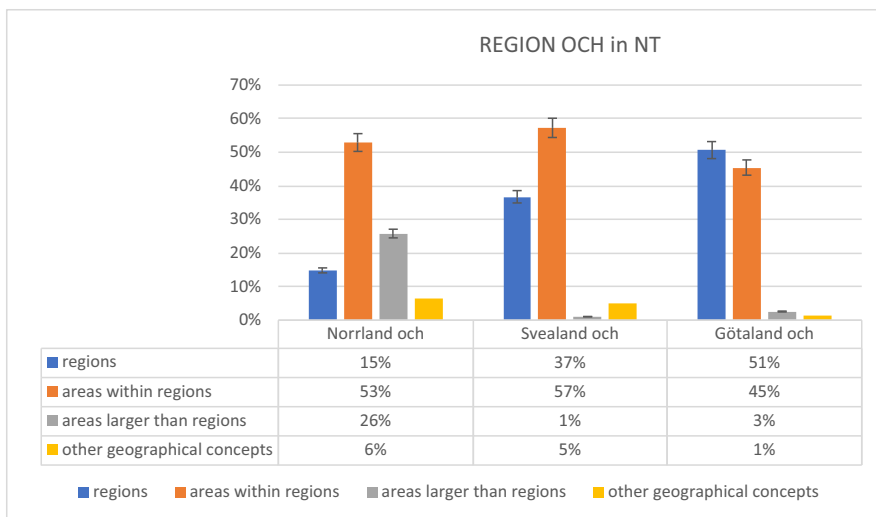


Figure 5. REGION OCH in news texts.

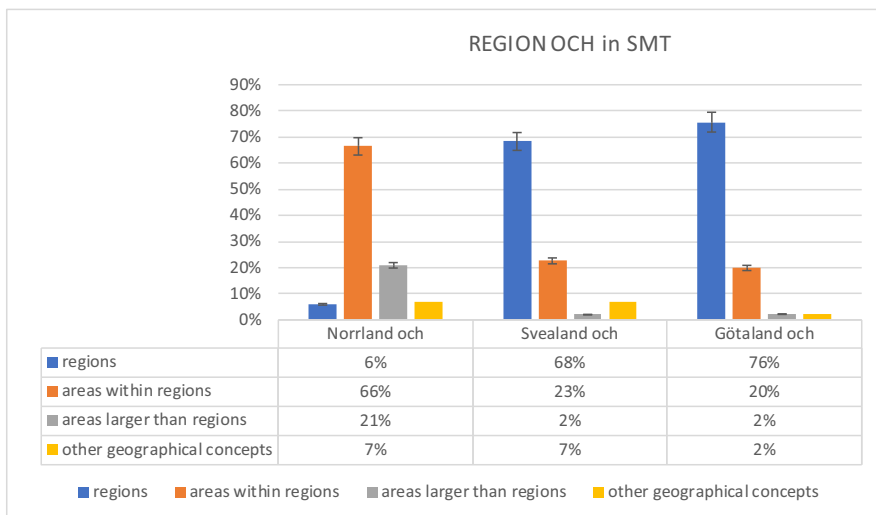


Figure 6. REGION OCH in social media texts.

5.1 Coordination with phrases for regions

Not only do the proportions of phrases for regions coordinated with *Norrland*, *Svealand*, or *Götaland* display differences between the regions, but so do the types of phrases coordinated with each of the regions and the order in which they occur.

Norrland is typically coordinated with *Svealand*, with *Svealand* commonly being the first of those coordinated, as in example (11).

- (11) Skogsgödslingen har stor betydelse för skogsproduktionen i
forest fertilization has big impact for forest production in
 Svealand och Norrland.
Svealand and Norrland.
 'Forest fertilization has a big impact on forest production in Svealand and
 Norrland.' (GT, *och Norrland*)

The proportion of *Norrland* instances coordinated with *Svealand* is the largest in NT (83% of both the *och Norrland* and *Norrland och* instances), somewhat smaller in GT (53% of *och Norrland* instances and 56% of *Norrland och* instances), and the smallest in SMT (35% of *och Norrland* and 38% of *Norrland och* instances). The *och Norrland* instances in all three subsets are coordinated with *Götaland* (12% in GT, 10% in NT, 5% in SMT) and those in SMT (13%) with *Norrland*. All these *och Norrland* instances coordinated with *Norrland* emphasize differences between two different referents of *Norrland*. As in (12), which reflects opposition to labeling Norrland as one homogenous unit, they reflect attitudes at odds with the asymmetry in categorization reflected in sentences (1)–(3) and the usage patterns of *Norrland*, *Svealand*, and *Götaland* observed in the present study (see Figures 1–6).

- (12) Ja, både Stockholm och Norrland bidrar, MEN det är
yes, both Stockholm and Norrland contribute, BUT there is
 lite skillnad på Norrland och Norrland
some difference on Norrland and Norrland.
 'Yes, both Stockholm and Norrland contribute, BUT there are differences
 between Norrland and Norrland.' (SMT *och Norrland*)

A relatively large proportion of the *Norrland* instances in GT and SMT are coordinated with both *Götaland* and *Svealand* in the order *Götaland*, *Svealand*, and *Norrland* (26% of instances in GT, 7% in NT, 38% in SMT), and the *och Norrland* instances in GT and SMT in the order *Svealand*, *Götaland*, and *Norrland* (9% in GT, 1% and 9% in SMT).

Most *och Svealand* instances (86% of instances in GT, 93% in NT, 89% in SMT) are coordinated with *Götaland*, with *Götaland* commonly being the first of the coordinated nouns as in (13).

- (13) Brandrisk-en är stor i Götaland och Svealand.
the fire risk is big in Götaland and Svealand.
 'The risk of fire is big in Götaland and Svealand.' (NT *och Svealand*)

Considerably fewer *och Svealand* instances (13% in GT, 7% in NT, 10% in SMT) are coordinated with *Norrland* or with combinations of all three regions.

The *Svealand och* instances are also commonly coordinated with *Götaland* (58% in GT, 59% in NT, 56% in SMT), but also with *Norrland* (33% in GT, 31% in NT, 21% in SMT) or with both *Norrland* and *Götaland* (6% of instances in GT, 8% in NT, 9% in SMT are part of the combination *Norrland*, *Svealand*, and *Götaland*, and 3% of them in GT, 2% in NT, and 14% in SMT are part of the combination *Götaland*, *Svealand*, and *Norrland*).

Götaland is typically coordinated with *Svealand* (84% of *och Götaland* instances in GT, 95% in NT, 88% in SMT, and 99% of *Götaland och* instances in GT, 92% in NT, 87% in SMT), with *Götaland* commonly being the first of the coordinated words.

To sum up, *Norrland*, *Svealand*, and *Götaland* differ not only in their tendencies to be coordinated with phrases for regions but also in terms of what regions they are coordinated with and in what order they are mentioned. Typically, the southernmost region is mentioned first, the northernmost last, and *Götaland* and *Svealand* coordinated.

Some of the differences observed are related to genre. Most examples of REGION and REGION coordination in NT concern weather reports. Because Sweden is an elongated country, the weather is usually reported in, on the one hand, the unit of *Götaland and Svealand* (the southern half of the country) and, on the other, in the unit *Norrland* (the northern half). Similarly, *Norrland* can be discerned as an economic unit, an agricultural unit, or a natural type in many examples in GT. The examples from NT and GT reflect *Norrland* as an established unit from many perspectives compared with other established units. Concerning at least weather, agriculture, and nature types, it makes sense to treat *Svealand and Götaland* as one unit rather than *Norrland and Götaland*, which is reflected in the many examples of *Svealand and Götaland*, the few examples of coordination between *Norrland and Götaland*, and the somewhat more frequent examples of coordination between the neighboring regions *Norrland and Svealand*.

5.2 Coordination with phrases for areas within regions

Svealand and *Götaland* instances are commonly coordinated with phrases for REGION-SPECIFIC areas (e.g. *södra Norrland* ‘the south of Norrland’). They comprise 22/28 *och Svealand* instances in GT (78%), 52/66 instances in NT (79%), and 14/40 (35%) in SMT, and 216/265 *Svealand och* instances in GT (82%), 222/268 in NT (83%), and 82/134 in SMT (61%). The *och Svealand* instances are coordinated with phrases for parts of *Götaland*, and the *Svealand och* instances with phrases for parts of *Norrland* or *Götaland*. In (14), *Svealand* is coordinated with a phrase for southern *Norrland*.

- (14) Under lördagen får vi sol över Svealand och södra Norrland.
under Saturday get we sun over Svealand and southern Norrland.
 ‘On Saturday we will get sunshine over Svealand and southern Norrland.’

The *och Götaland* instances coordinated with phrases for REGION-SPECIFIC areas comprise 12/18 instances in GT (67%), 21/30 instances in NT (70%), and 9/37 (24%) in SMT, and the *Götaland och* instances 18/30 of instances in GT (60%), 25/35 in NT (71%), and 19/36 in SMT (53%). All these instances but one (coordinated with a phrase for a more specific part of *Götaland*) are coordinated with phrases for parts of *Svealand*. They are typically from, or related to, weather forecasts and involve topics such as temperatures, cloudiness, precipitation, wind, seasons, traffic situations, water yield, and wildfires in the areas, but there are also topics about forestry, employment, electricity prices in the areas, or the spread of animals such as

wild boars, mites, or spruce bark within them. One example is (15), which deals with an expected increase in the proportion of spruce forests in *Götaland* and *the southern part of Svealand*.

- (15) Andel-en granskog förväntas öka i Götaland och
the proportion spruce forest is expected increase in Götaland and
 södra del-en av Svealand.
the southern part of Svealand.

'The proportion of spruce forests is expected to increase in Götaland and the southern part of Svealand.'

Considerably fewer of the *Svealand* and *Götaland* instances coordinated with phrases for areas within regions involve phrases for PROVINCES (e.g. *Skåne*), PROVINCE-SPECIFIC categories (e.g. *Gävleborgs kustland* 'the Gävleborg coast'), vague AREAS (e.g. *Norrlandskusten* 'the Norrland coast'), or PLACES (e.g. *Norrköping*). The most common combinations include *Götaland* instances in all three subsets (3/18 *och Götaland* instances in GT [17%], 6/30 in NT [20%], 20/37 in SMT [54%], and 5/30 *Götaland och* instances in GT [17%], 9/35 in NT [26%], 12/36 in SMT [33%]) and *och Svealand* instances in SMT (17/40 [(42%)] and NT (10/66 [15%]), and *Svealand och* instances in SMT (25/134 [19%]). The *Svealand* instances are primarily coordinated with provinces in *Götaland* or *Norrland* (13/17 instances), but the *Götaland* instances with provinces within *Götaland* (15/20 of *och Götaland* instances in SMT, 4/6 in NT, and 8/12 *Götaland och* instances in SMT) and particularly with *Skåne*, *Gotland*, or *Småland*.

- (16) Snart är vi ett helt land, bara skåne [sic] och Götaland kvar.
soon are we a whole country, only skåne and Götaland left.
 'Soon we will be a whole country, only skåne and Götaland left.'
 (SMT, *och Götaland*)

The *Götaland* instances specify the location of the province within this region (i.e. describe a part-whole relation) rather than compare *Götaland* to a province denoted by the coordinated word.

Relatively many *Götaland* instances in GT are coordinated with phrases for VAGUELY DELIMITED AREAS within regions (e.g. *Mälardalen* in *Svealand*). They make up 3/18 of *och Götaland* instances (17%) and 7/30 *Götaland och* (23%) instances. Less than 14% of *Svealand* instances and *Götaland* instances are coordinated with phrases for AREA-SPECIFIC categories.

Unlike the *Svealand* and *Götaland* instances, the *Norrland* instances are rarely coordinated with phrases for REGION-SPECIFIC instances. They comprise 86/517 *och Norrland* instances in GT (17%), 24/183 instances in NT (13%), and 4/473 (1%) in SMT, and 19/219 *Norrland och* instances in GT (9%), 40/164 instances in NT (24%), and 5/89 in SMT (6%). Instead, most *Norrland* instances are coordinated with phrases for PROVINCES, and particularly with the neighboring province of *Dalarna* in northern *Svealand*, the island of *Gotland*, and the southernmost province of *Skåne* showing that the asymmetry reflected in sentence (3) is common. Instances coordinated with PROVINCES include 286/517 *och Norrland* instances in GT (55%),

74/183 in NT (40%), and 267/473 in SMT (56%), and 122/219 *Norrland och* instances in GT (56%), 60/164 in NT (37%), and 40/89 in SMT (45%).

Moreover, the *och Norrland* instances in NT (66/183 instances [37%]) and SMT (162/473 [34%]) and the *Norrland och* instances in NT (34/164 instances [21%]) and SMT (33/89 [37%]) are commonly used with phrases for places such as *Malmö* in sentence (1), *Göteborg, Sundsvall, Falun-Mora* in sentence (2), and *Stockholm* and *Göteborg* in sentence (3). This type of coordination is unusual in instances coordinated with *Svealand* or *Götaland* (0–14% of instances), further highlighting the asymmetry in categorization reflected in sentences (1)–(2).

Norrland instances in GT are commonly used with phrases for more vaguely defined AREAS within regions, e.g. *Bergslagen* and *Mälardalen* in *Svealand* and *Västkusten* (the Swedish ‘West Coast’) in *Götaland*. They make up 90/517 of *och Norrland* instances (17%) and 43/219 *Norrland och* instances (20%) in this subset. Coordination with these phrases typically builds on similarity and properties shared with *Norrland*. For instance, *Norrland* and *Bergslagen* share properties related to social issues, the economy, and natural resources.

A small proportion of the rest of *Norrland* instances coordinated with phrases for places within regions denote AREAS ($\leq 12\%$), PLACES ($\leq 9\%$), OR PROVINCE-SPECIFIC ($\leq 5\%$), AREA-SPECIFIC ($\leq 1\%$), or PLACE-SPECIFIC ($\leq 3\%$) categories.

In conclusion, *Svealand* and *Götaland* are more commonly coordinated with REGION-SPECIFIC phrases than *Norrland*. Instances such as these resemble REGION AND REGION coordination in the sense that the region is involved. The noun for the region is part of the coordinated phrase and its function is to provide information on which part of the region is relevant. Coordination with *Norrland* is more heterogeneous. It primarily involves phrases for PROVINCES but also phrases for areas, places, or more specific parts of them. The differences align with the previously observed asymmetry in categorization involving *Norrland*, on the one hand, and *Svealand* and *Götaland*, on the one hand.

5.3 Coordination with phrases for areas larger than regions

Most instances coordinated with phrases for areas larger than regions include COUNTRY-SPECIFIC areas. They include both *och Svealand* and *och Götaland* instances in NT and *Svealand och* and *Götaland och* instances in GT, as well as all six *Svealand och* instances and all three *Götaland och* instances in NT. The *Svealand* and *Götaland* instances are primarily coordinated with phrases for the south of Sweden (e.g. *sydsverige* and *södra Sverige*), sometimes including both these regions in their references and sometimes suggesting that the south of Sweden is south of *Svealand*. In (17), one of the seasons is coming to *Svealand* and to parts of southern Sweden.

- (17) Nu kommer den meteorologiskt in även i Svealand och delar
now comes it meteorologically in also in Svealand and parts
 av södra Sverige.
of southern Sweden.

‘Now it meteorologically arrives also in *Svealand* and parts of southern Sweden.’ (NT, *Svealand och*)

The *Norrland* instances are primarily coordinated with phrases for mid-Sweden (e.g. *Mellansverige* and *Mellersta Sverige*), indicating that this part of the country is south of Norrland.

- (18) I mellersta Sverige och Norrland kommer det vara fint
in mid Sweden and Norrland comes it be fine
 sommarväder i dag.
summer weather today.
 'In mid Sweden and Norrland there will be fine summer weather
 today.' (NT, *och Norrland*)

Considering that Sweden's geographical center is in the south of Norrland, instances such as these suggest that the north, south, and middle of Sweden are defined by population distribution (its demographic center) rather than geography.

Most *och Götaland* and *Svealand och* instances in SMT (3/4 instances of each subset) are also primarily coordinated with COUNTRY-SPECIFIC areas, but the rest of the instances are coordinated with phrases for COUNTRIES. The latter group includes *och Norrland* (15/28 instances [54%]), *Norrland och* (59/95 [62%]), and *och Svealand* instances (9/12 [75%]). The relatively large proportion of *Norrland* instances coordinated with COUNTRIES align with the somewhat sweeping categorizations of *Norrland* reflected in instances (1)–(3). Sentence (19) shows that asymmetry in categorization may be grounded in the identity of speakers.

- (19) Skåning som jag är trodde jag också att Norrland var
skåning as I am thought I also that Norrland was
 samma sak som Norge, trots att jag faktiskt hade
same thing as Norway, although that I actually had
 varit i både Norrland och Norge.
been in both Norrland and Norge.
 'Being from Skåne, I thought Norrland was the same as Norway, although
 I had been to both Norrland and Norway.' (SMT, *Norrland och*)

5.4 Coordination with phrases about other geographical units

Instances combined with phrases for DIRECTIONS (e.g. *norrut* 'north', *uppåt* 'up', *söderut* 'south') are primarily found in the *Svealand och* subsets and about the weather or related phenomena in all three subsets. Most of them (12/17 instances in GT, 18/22 in NT, 28/35 in SMT) refer to a northward direction as in (20), again highlighting less specific and more sweeping information on areas further north.

- (20) I skrivande stund är det nu snötäckt från Svealand
in writing moment it is now snow covered from Svealand
 och norrut (skåne [*sic*] och Blekinge har också snö)
and north (Skåne and Blekinge have also snow)
 annars är det barmark.
otherwise it is bare ground
 ‘At the time of writing, the ground is covered by snow from Svealand
 and north (there is snow in Skåne and Blekinge); elsewhere it is free
 of snow.’ (SMT, *Svealand och*)

Instances coordinated with phrases for vague areas such as *glesbygd* (‘sparsely populated area’) *landsbygd* (‘rural area’), *skogslän* (‘forest county’), *inlandet* (‘the inland area’) tend to be found in the *Norrland* instances in GT (30/30 *och Norrland* instances for other geographical areas in this subset, and 103/110 *Norrland och* instances), the *och Norrland* instances in SMT (36/38 instances) and the *Norrland och* instances in NT (13/20 instances). This type of coordination typically highlights properties that the areas (are perceived to) have in common with *Norrland*.

- (21) Billiga hus hittar man på land-et och Norrland
cheap houses find one on the countryside and Norrland
 medan de dyrare hittas i villaområdena i städerna.
whereas the expensive are found in the house areas in the suburbs
 ‘Cheap houses may be found in the countryside and in *Norrland* while the most
 expensive ones are found in residential neighborhoods in the suburbs.’
 (SMT, *och Norrland*)

6. Concluding discussion

Results show that the *Norrland*, *Svealand*, and *Götaland* regions are far from treated as three equal entities of the same category. *Svealand* and *Götaland* are more commonly coordinated with nouns for regions than *Norrland* in all subsets (RQ1), and *Norrland* is more commonly coordinated with phrases for geographical places within the regions of Sweden compared with *Svealand* and *Götaland* in most subsets (RQ2). This suggests that the asymmetry in categorization reflected in sentences (1)–(3) is widespread in all three types of discourse investigated here, ranging from texts in the most private and presumably least (politically) correct discourse SMT to the public discourse NT to official GT.

We may conclude that categorization strategies involving *Norrland* are notably different from those involving *Svealand* and *Götaland*. Not only are towns, provinces, and other areas within *Norrland* labeled as *Norrland* but *Norrland* is also compared with even larger geographical categories comprising countries and continents more commonly than *Svealand* and *Götaland*.

The results align with those obtained by Eriksson (2010) in her analysis of representations of *Norrland* in contemporary news texts. Representations of *Norrland* tend to come with less detail than those of other regions. At the same time, they are accompanied by certain specific connotations. The patterns are seen not only in SMT but also in GT and NT. Considering the similarities in the proportions

of phrases for regions, areas within regions, areas that are larger than regions, and other geographical categories in all three OCH REGION subsets (Figures 1–3) and SMT in the REGION OCH subsets, it appears safe to say that they are highly representative of coordination with *Norrland*, *Svealand*, or *Götaland*. Moreover, considering the similarities between the subsets, they do not appear due to how the subcorpora were compiled.

However, the REGION OCH subsets in GT and NT (Figures 5 and 6) are partly different due to the large proportions of REGION-SPECIFIC coordination with *Svealand* in official contexts about the weather or the spread of animals and vegetation in parts of other regions. Presumably, the high frequency of the *Svealand* instances coordinated with these phrases is related to the topics they address and the size and location of this region. Given that most instances deal with the weather and the spread of animals or vegetation, which may be similar in adjacent regions, there may be numerous reasons for thinking and talking about *Svealand* and parts of adjacent regions this way. Still, the *Svealand* and *Götaland* instances coordinated with phrases for AREAS WITHIN REGIONS are qualitatively different from those coordinated with *Norrland*. The *Svealand och* instances include phrases for areas within the other two regions (e.g. *södra Norrland* ‘south Norrland’), the *Götaland* instances include phrases for REGION-SPECIFIC areas within Götaland specifying a part of this very region, and the *Norrland* instances include phrases for PROVINCES, AREAS, or PLACES in Norrland, again accentuating the asymmetry in categorization strategies. As in (1), where *Norrland* is used for the place from which a representative comes, these *Norrland* instances commonly reflect sorting places within Norrland into the considerably larger category of *Norrland*. In conclusion, categorization strategies involving Norrland are notably different from those involving *Svealand* and *Götaland*. Not only are towns, provinces, and other areas within Norrland labeled as *Norrland*, but Norrland is also compared with even larger geographical categories comprising countries and continents more commonly than *Svealand* and *Götaland*. In contrast, the *Svealand* instances coordinated with phrases for REGION-SPECIFIC areas do not necessarily result from asymmetric categorization. Given that the noun for the other two regions is typically part of the coordinated phrase, speakers appear to have had the regions in mind when using these phrases but then qualified what specific parts of them are relevant.

There may be several reasons for the differences in categorization reflected in the present study. Some might be related to Sweden’s geography, population density, and categorization processes grounded in speakers’ body–world knowledge (Lakoff 1987). Considering the length and elongated shape of Sweden and the fact that most Swedes live in either *Svealand* or *Götaland*, a less detailed categorization of geographical entities within Norrland might be close at hand for most speakers of Swedish even though Norrland, *Svealand*, and *Götaland* are all regions. Put differently, they might be because areas in the north are further away and hence less familiar to most speakers of Swedish than areas in the south. The relevance of the identity of the speakers is similarly suggested by the order in which the regions and their parts are mentioned – for example from the south to the north or in connection with REGION-SPECIFIC coordination from *Svealand* in the middle to part of *Götaland* in the south or *Norrland* in the north.

Yet other differences are related to genre. For instance, there were differences between subsets related to the fact that different genres serve distinct functions, have different target audiences, and focus on different perspectives of the regions, and possibly also to differences between the denotations of *Norrland*, *Svealand*, and *Götaland* in the subcorpora. In GT, *Norrland* is either used for the five counties of Norrland, the forest-rich counties in Sweden (including some counties in Svealand), or an area used in international comparisons (national region). No matter if each of these official uses is precise, the fact that the areas denoted by *Norrland* vary in a standardized way in official contexts such as these contributes to the conceptualization of *Norrland* as a unit with low specificity/high degree of generality.

In contrast, instances in SMT carry a distinct personal flavor. They are often connected with local issues, and here the concept of *Norrland* is more intricately connected with the provinces of *Norrland*. In SMT, *Norrland* appears synonymous with the north, as shown by example (22), taken from a context discussing people's understandings of which landscape fits which region. Here, the subjective perspective is highlighted.

- (22) Ser du skillnader på norrland och norrland? [sic] För mig
 see you differences on *Norrland* and *Norrland*. For me,
 är *Norrland* norr om mig själv.
 is *Norrland* north of myself.

'Do you see differences between *norrland* and *norrland*. For me, *Norrland* is north of myself.' (SMT *och Norrland*).

Genre differences such as these may thus have contributed to the differences between the subsets.

The asymmetries observed in this data have implications. Given that category labels induce focus on commonalities (Althaus & Plunkett 2016), uses of the noun for the region *Norrland* for its parts (e.g. its towns and provinces) highlight commonalities between the parts denoted (no matter what they are like, they are all like 'Norrland') and hide attributes specific to each more specific place. As observed by Eriksson (2010:57), they involve using more general aspects of the region for specific counties, villages, and towns, whereby what is universal represents what is specific, and what is specific to these places is overlooked. In that way, uses such as these may reinforce stereotypes (everything in *Norrland* is one way, whereas the parts of *Svealand* and *Götaland* differ in substantial ways).

The present study was based on a mix of quantitative and qualitative analysis of no less than 18 subsets, each comprising up to 1,000 *Norrland*, *Svealand*, or *Götaland* instances. Still, it only scratches the surface of how these nouns are used. Qualitative semantic analyses of this data type are potentially fruitful complements to the present study. Considering the subjective quality of SMT, in-depth investigations of SMT might be particularly fruitful. Furthermore, semantic analyses of the phrases for the non-geographical categories (e.g. *snö* and *Norrland* 'snow and *Norrland*') removed from the study potentially show what semantic fields are associated with the three regions and what connotations the nouns for the regions have. Second, although the results of this study are representative of coordination with *Norrland*, *Svealand*, or *Götaland* in GT, NT, and SMT, there is more to learn

about the regions from other constructions, including these nouns (e.g. those with prepositions signaling movement towards these areas, or with adverbs of place).

Finally, it may be fruitful to investigate how the asymmetric categorization of place observed here relates to the attitudes of speakers and how it is perceived among those living in places referred to as a region, and the region, in turn, portrayed as a homogenous unit with low specificity/high degree of generality. Recent debates in social and established media show that it has brought about a great deal of frustration that may only sometimes be understood by those categorizing the Swedish regions in asymmetric ways. Yet another question concerns whether the asymmetry observed here is a Swedish phenomenon or whether it may be observed in language about areas in other Nordic countries and the rest of the world.

Acknowledgments. We express our gratitude to three anonymous reviewers for their constructive feedback, which has greatly improved the quality of this paper.

Notes

1 List of abbreviations: GT = governmental texts; NT = news texts; SMT = social media texts; RQ = research question; Sw. = Swedish.

2 Proper names may be seen as kinds of nouns or treated as a part of speech separate from nouns (Teleman et al. 1999:117). We refer to geographical names as nouns throughout.

3 Urban areas are here defined as areas with more than 200 inhabitants.

4 The majority of these *och Norrland* instances in all three subcorpora (49% in NT, 60% in MT, 72% in SMT), *och Svealand* instances in MT (83%) and SMT (68%), and *och Götaland* instances in all three subcorpora (60% in NT, 78% in MT, 85% in SMT) were coordinated with phrases for the same type of geographical location as the first phrase. However, the majority of these *och Svealand* instances (50%) in NT were coordinated with phrases for several different geographical locations. Instances that were not coordinated with phrases from either of these categories were coordinated with ones for smaller or larger geographical areas or (two of them) with nouns for non-geographical entities (as in e.g. *Uppsala, musik* ['music'] *och Norrland*).

References

- Ainiala, Terhi & Ola Östman (eds.). 2017. *Socio-onomastics: The pragmatics of names*. Amsterdam: John Benjamins.
- Althaus, Nadja & Kim Plunkett. 2016. Categorization in infancy: Labeling induces a persisting focus on commonalities. *Developmental Science* 19(5). 770–780.
- Asker, Björn. 2009. *Hur riket styrdes: Förvaltning, politik och arkiv 1520–1920* [How the kingdom was governed: Administration, politics and archives 1520–1920], 2nd edn. Stockholm: Riksarkivet.
- Borin, Lars, Markus Forsberg & Johan Roxendal. 2012. Korp: The corpus infrastructure of Språkbanken. In Markus Forsberg & Johan Roxendal (eds.), *Proceedings of LREC 2012*, 474–478. ELRA.
- Cruse, D. Alan. 1977. The pragmatics of lexical specificity. *Journal of Linguistics* 13(2). 153–164.
- Edlund, Lars-Erik & Tore Frängsmyr (eds.). 1995. *Norrländsk uppslagsbok: Ett uppslagsverk på vetenskaplig grund om den norrländska regionen* [Norrländsk encyclopedia: A scientific reference work on the Norrland region], vol. 3. Umeå: Norrlands universitetsförlag.
- Edlund, Lars-Erik & Anna K. Greggas (eds.). 2000. *Kulturgräns norr: Studier av kulturella förändringsprocesser på nordskandinaviskt område* [Cultural border north: Studies of cultural reformation processes in northern Scandinavia]. Umeå: Umeå University.
- Eriksson, Madeleine. 2010. *(Re)producing a periphery: Popular representations of the Swedish north*. Umeå: Umeå University.

- Förordning (EU) 2019/2088 av den 27 november 2019 om hållbarhetsrelaterade upplysningar som ska lämnas inom den finansiella tjänstesektorn [Regulation (EU) 2019/2088 of 27 November 2019 on sustainability-related information to be provided within the financial services sector]. 2019. https://eur-lex.europa.eu/legal-content/SV/TXT/?uri=uriserv%3AOJ.L_.2019.270.01.0001.01.SWE&to c=OJ%3AL%3A2019%3A270%3ATOC
- Helleland, Botlov. 2012. Place names and identities. In Botlov Helleland, Christian-Emil Ore & Solveig Wikström (eds.) *Names and identities* (Oslo Studies in Language 42), 95–116. Oslo: University of Oslo.
- Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lupyan, Gary, David. H. Rakison & James L. McClelland. 2007. Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science* 18(12). 1077–1083.
- Mark, David M., Barry Smith & Barbara Tversky. 1999. Ontology and geographic objects: An empirical study of cognitive categorization. In Christian Freksa & David M. Mark (eds.), *Spatial information theory: Cognitive and computational foundations of geographic information science (COSIT'99)*, 283–298. Berlin/Heidelberg: Springer.
- Medin, Douglas L. & Marguerite M. Schaffer. 1978. Context theory of classification learning. *Psychological Review* 85(3). 207–238.
- Robinson, Christopher W. & Vladimir M. Sloutsky. 2007. Linguistic labels and categorization in infancy: Do labels facilitate or hinder? *Infancy* 11(3). 233–253.
- Rosch, Eleanor. 1973. Natural categories. *Cognitive Psychology* 4(3). 328–350.
- Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch & Barbara Bloom Lloyd (eds.), *Cognition and categorization*, 27–48. Hillsdale, NJ: Lawrence Erlbaum.
- Rune, Olof. 1995. Limes Norrlandicus. In Lars-Erik Edlund & Tore Frängsmyr (eds.), *Norrländsk uppslagsbok: Ett uppslagsverk på vetenskaplig grund om den norrländska regionen* [Norrländsk encyclopedia: A scientific reference work on the Norrland region], vol. 3. Umeå: Norrlands universitetsförlag.
- Sörlin, Sverker. 2023. *Framtidslandet* [The land of the future], 2nd edn. Luleå: Teg Publishing.
- Teleman, Ulf, Staffan Hellberg & Erik Andersson. 1999. *Svenska Akademiens grammatik*, vol. 2, *Ord* [Swedish Academy grammar, vol. 2, Words]. Stockholm: Svenska Akademien.