

Annotation scheme for metonymies (AS1)

Katja Markert
School of Computing
University of Leeds
Woodhouse Lane
Leeds LS2 9JT
UK

`markert@comp.leeds.ac.uk`

Malvina Nissim
Language Technology Group
Division of Informatics
2 Buccleuch Place
Edinburgh EH8 9LW
UK

`markert@comp.leeds.ac.uk`

June 7, 2005

Chapter 1

Introduction and Purpose

In this chapter we introduce the purpose of the annotation. In addition we present our underlying assumptions and an outline of the annotation procedure.

1.1 Introduction

In metonymy one entity is used to refer to another related entity. A typical and often-cited example is

(1.1) “*The ham sandwich is waiting for his check.*”

Here it is not the ham sandwich (A) that is waiting, but the person (B) that is eating the ham sandwich. Therefore we say A stands for B. In Example (1.1) it is not the city that sleeps, but (most of) the people living in the city.

(1.2) “*Paris sleeps*”

Example (1.1) is a fairly innovative and unconventional metonymy, which occurs only in special contexts; Example (1.2) is conventional — it can be applied to all cities or countries. We call these latter metonymies *schematic* and talk about A-for-B metonymies. In the above case we talk about a Place-for-People metonymy or a Place-for-People metonymic *pattern*.

In contrast to metaphors the two related items are not related by similarity; so in the above examples the “ham sandwich” and the customer are not similar, but related by an eating relation. An example of metaphor would be calling a person “a lion” because he is strong as a lion (similarity).¹

This annotation scheme handles both innovative and conventional metonymies. We will only provide explicit, predefined tags for schematic metonymies. However, in case both the literal as well as the predefined metonymic patterns do not apply, an unconventional metonymic or a metaphoric reading might be intended, which will be annotated with a single tag each. We will also annotate mixed readings where both a literal and a metonymic reading is invoked.

1.2 Purpose, Coverage and Assumptions

1.2.1 Purpose

Metonymy recognition and partial interpretation.

¹Obviously the distinction between metaphor and metonymy is not always that clear. Thus, in an Organisation-for-Members metonymy, one could argue that the organisation is also similar to a person.

1.2.2 Material and Coverage

This annotation scheme intends to cover both common nouns (e.g., “country”), proper names (e.g., “Russia”) and adjectives derived from them (e.g., “Russian”). For identification and classification of proper names as well as identification of their extent we follow mainly the MUC-7 guidelines.

The guidelines have been developed and tested using text excerpts from the British National Corpus (BNC), containing a proper name or common noun of a given semantic class, e.g., author names or locations. The noun to be annotated is from now on called Possible Metonymic Word (PMW). The text excerpts contained 2 sentences before PMW, the current sentence and 1 sentence after the PMW. The PMW can be a multi-word item, like “William Shakespeare” (see the MUC-7 guidelines). In the examples in this document, assume that the noun in bold face is the PMW to be annotated.

An example of such a text excerpt is:

*“Representatives of the Group of 24 (G-24) industrialised countries, which co-ordinated aid for eastern Europe and former Soviet Asia, met in the Albanian capital, Tirana, on July 22 in order to discuss with the Albanian authorities and international financial institutions how aid to Albania might be accelerated. The G-24 group expressed readiness to provide **Albania** with food aid until the 1993 harvest, and beyond if necessary. Italy was a major donor among European countries and Greece announced that it would make available US 70,000,000 for agricultural development, which was given high priority.”*

1.2.3 Assumptions

The input to the metonymy annotation are texts where entities of particular semantic classes are already marked up with respect to their extent and their general semantic class (*base type*). Thus the input to the metonymy annotation (in an XML format) looks like in Table 1.1:

<LOCATION> Hungary </LOCATION> took similar actions . . .

Figure 1.1: XML input example for the semantic class LOCATION

Texts that are marked up in this fashion already exist for proper names: As an example MUC-7 texts are already marked for location, person and organisation proper names that can then be analysed for metonymies. However, such texts do not normally exist for common nouns. Therefore we have developed specific guidelines for identifying these base types (see AS2).

We analysed excerpts from the BNC, identified base type entities in these texts (using similar guidelines to MUC-7 for proper names and AS2 for common nouns) and then analysed these in a second step for metonymic usage.

We develop sets of annotation schemes for separate semantic classes as they determine to a large degree which metonymies a word can undergo. Thus, e.g. location names frequently undergo **Place-for-People** or **Place-for-Event** metonymies. Nevertheless, we provide some class-independent metonymic patterns relevant for all semantic classes as well as the option to annotate metonymies that do not fall into the patterns provided (unconventional metonymies) and to annotate mixed readings (readings where two metonymic patterns or a metonymic pattern and the literal reading are invoked at the same time). The structure of the annotation scheme therefore looks like in Figure 1.2.

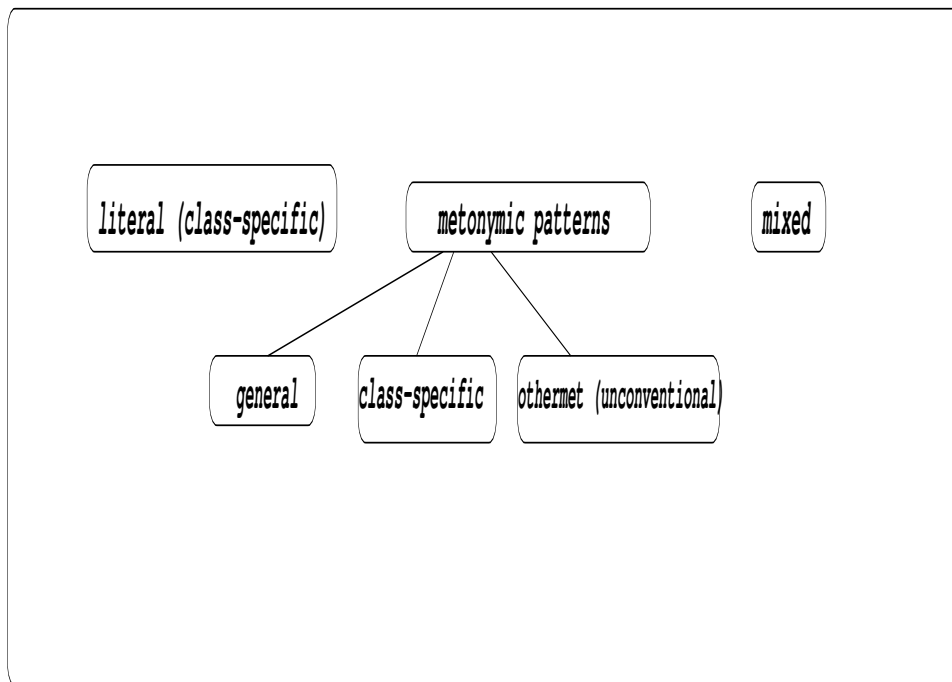


Figure 1.2: The structure of the annotation scheme

The output of the metonymy annotation is shown in Figure 1.3, identifying a metonymic reading of the kind *Place-for-People* with the additional specification that the people meant are officials of the state Hungary.

```

<LOCATION      reading="metonymic"      metotype="place-for-people"
subtype="OFF"> Hungary </LOCATION> took similar actions ...
  
```

Figure 1.3: XML output of metonymy annotation for the class `LOCATION`

The annotation scheme will be successively extended to more semantic classes.

1.3 Annotation Procedure

Please follow the following steps in annotation.

1. Please read the full text given before you decide on readings and make sure you understand it.
2. The annotation is performed using the MATE annotation tool <http://mate.nis.sdu.dk/>. It already gives you the choice of values for all attributes mentioned. We do not describe the annotation tool in this document.
3. Your main task is to assign a literal or metonymic reading to the PMW, which is already marked for its base type (e.g. `LOCATION`, `PERSON` or `ORGANISATION`) according to the MUC-7 guidelines for proper names or according to the guidelines defined in AS2 for common nouns.
4. The guidelines are given separately for different classes of nouns (e.g. `LOCATION` or `PERSONS`). For every metonymic pattern as well as for the literal

reading we give a **replacement test** and/or other guidelines. Replace the PMW with the term given in the replacement test in the current utterance; then judge whether the ensuing reading is the one intended in the current context. If so, the replacement test is *successful*, otherwise it is not successful. As in some cases the replacement tests are not sufficient, we have also accumulated other guidelines.

5. Then choose the reading indicated by the test/set of guidelines that is successful (either as the value “literal” for the attribute “reading” or as “metonymic” for the attribute “reading” plus the specific metonymic pattern in the attribute “metotype”). **If two tests are successful and one of them indicates the literal reading, then give preference to the literal reading!**
6. If no replacement test is successful, one of the following cases might hold:
 - It might be a *mixed reading*, meaning that an appositive or conjunction or several modifications seem to call for different readings (see Section 2.2). Mark these cases as “mixed” for the attribute “reading” and give the two contributing readings as the value for the attribute “mixedtype”.
 - It might be that a metaphor is present — mark these readings as “other” for the attribute “reading”.
 - It might be that a metonymy is present that does not fall under the specified categories. Mark these as the value “metonymic” for the attribute “reading” and then as “othermet” for the attribute “metotype”.
7. Sometimes the metonymic pattern is a pattern with several subordinate patterns, e.g. `Place-for-People` is divided into `Place-for-Organisation`, `Place-for-Population` etc. In these cases, decide on the general pattern first. If you are sure about the subordinate pattern, mark it as the value of the “subtype” attribute.
8. You can mark any additional comments in the attribute “Notes”.

1.3.1 Structure of this document

This document is structured as follows. In the next Chapter, the class-independent metonymic patterns (Section 2.1) as well as the requirements for mixed readings (Section 2.2) are described. Each of the following chapter contains the guidelines for a specific semantic class.

Chapter 2

Class-independent readings

Class-independent readings include class-independent metonymic patterns (Section 2.1) and mixed readings (Section 2.2).

2.1 Class-Independent Metonymic Patterns

There are two metonymic patterns that can be applied to all concrete nouns.

2.1.1 Object-for-Name

A metonymic pattern which can be used for *all* nouns is **Object-for-Name**. It means that the noun *x* is not used to refer to an object (or a collection of objects etc.), but as a mere string to indicate the name of the object.

Test. The replacement test to be carried out is to replace *x* by “(by) the name *x*”. This should be the first test carried out in every instance.

Examples.

1. “*This is called a **printer***”.
2. “*I like the name **Shakespeare***”

Remarks and Hints. Although often indicated by collocations such as “called” or “name” this need not always be the case. So “*I like **David***” can be used to indicate “*I like the name **David***” in a suitable context.

Possible confusions. Common nouns, differently from proper names, can be used generically, and this has some consequences on replacement tests. In “*he didn’t write his name but only the **town***”, say on the back of an envelope, an **object-for-name** metonymy would require an interpretation where the person wrote the word **town**. The most likely interpretation, though, is one where what is written on the envelope is actually the name of a *specific* town (e.g. “Brighton”), rather than ‘town’. In this case we opt for an **object-for-representation** pattern, since it is indeed one of the possible representations of the specific town.

Marking. Specify the value “metonymic” for the attribute “reading” and specify the value **Object-for-Name** for the attribute “metotype”.

2.1.2 Object-for-Representation

A metonymic pattern which can be used for *all* (concrete) nouns is **Object-for-Representation**. The noun *x* is not used to refer to the actual object (or a collection of objects etc.), but to refer to a representation of the object, e.g. on a photo, a map or a painting. For an organisation we regard the logo of the organisation as its representation.

Test. The replacement test to be carried out is to replace *x* by “representation (photo, painting, drawing etc.) of *x*”. This should be the second test carried out for every instance.

Examples.

1. “*This is **Shakespeare***” (pointing to a picture of Shakespeare).
2. “***Malta** is here.*” (looking at a map).
3. “*In his film, he shows a different **Paris**. His **Montmartre** is...*”

Remarks and Hints. Note that this metonymy can often only be discovered in a context comprising several utterances. The kind of representation medium does not matter, e.g. on coins, on photos, on the web, in a sculpture...

Possible confusions. Possible confusion with the literal meaning in utterances like

- “*This is a picture of **Shakespeare***”
- “*I have painted **Shakespeare**.*”

In the first example, the occurrence of the word is **literal** as the representation is given explicitly (note the replacement test does not work: “This is a picture of the representation of Shakespeare” has a different meaning.) In the second example, both replacement tests (literal “I have painted the person Shakespeare” and metonymic “I have painted a representation of Shakespeare”) are successful. In these cases we decide on the literal reading.

Marking. Specify the value “metonymic” for the attribute “reading” and mark the value **Object-for-Representation** for the attribute “metotype”.

2.2 Mixed readings

A mixed reading can only occur if the expression takes part in two predication contexts, e.g., it is the subject of a verb as well as having an adjective modifier, imposing different constraints on the reading. It is therefore similar to a zeugma. What is important is that the annotator must be able to give the two predications that induce the mixed reading explicitly before assigning a mixed reading. (Exceptions regarding appositions and lists are noted below). Possibilities are:

- **Two modifiers.** Two modifiers of the PMW calling for different restrictions, including adjectival modifiers as well as prepositional post-modifiers, genitive modifiers, appositions and relative clause modifiers. An example is: “*a coastal **state** that is very afraid of rising sea-water levels*”. The modifier “coastal” seems to point to a literal reading whereas the modifier “that is very afraid...” can only apply to people.

- **A head and a modifier.** The head of the noun phrase and a modifier giving different constraints. Heads include another noun phrase in a genitive or prepositional construction or a verb. An example is: “*The conference of 8 Arctic seaboard **countries***” where “arctic seaboard” points to a literal reading of “countries” whereas “conference of” points to a Place-for-people reading.
- **Appositions and lists.** If the proper name N occurs (as a modifier) in an apposition or a list, both the restriction coming from the head H of the apposition or list as well as any restrictions applying to that head H are taken into account. Examples make this clear: “*a conference of the 8 Arctic seaboard countries — **Sweden, Norway** – took place in 1984.*”. If Sweden is to be annotated, we also take all the restrictions of the head of the listing “countries” into account. Therefore we annotate a mixed reading for Sweden. The same holds for appositions not introduced by dashes, but by blanks, commas or the words “namely”, “e.g.”, and “i.e.”. Another example is “*the countries **Afghanistan and Pakistan** did sign the treaty.*”, leading to a mixed reading. If the lists include an additional “of” as in “*the country of **Sweden** did not sign the treaty*” this does not make any difference, i.e. the reading is still mixed. Accordingly, then in “*the **country** of Sweden did not sign the treaty*”, **country** has to be annotated as mixed as well, given that “Sweden” here triggers the literal reading.
- **Caution: “like” and “such as”.** For these comparisons we do not regard the restrictions coming from the head H of the name. So, in “*Countries like **Afghanistan** signed the treaty.*”, “Afghanistan” is to be annotated as literal only. The reason for this is that the comparisons operate on “countries” only, not involving the government.
- **Head verb + gerund or Head verb + infinitive.** If the name N has a verb V as a head followed by a gerund or an infinitive, we take both predicates into account. An example is “***Sweden** claims to be a great power.*”. This leads to a mixed reading as “claim” is a human action whereas “being a great power” refers to the state.
- Other restrictions, especially coreference, are not taken into account.

Chapter 3

Class-specific guidelines

3.1 The semantic class LOCATION

At the moment, the annotation scheme for LOCATION covers the proper names of countries, cities, villages and towns, administrative districts like American states, Italian administrative regions or German *laender* (corresponding to the proper names in WordNet under “district”) as well as seas, lakes, rivers, deserts, stellar bodies, mountains and valleys as well as location common nouns. We follow the MUC-7 annotation guidelines for determining the extent of a location name and the guidelines specified in AS2 for location common nouns.

The guidelines cover the corresponding adjectives derived from the proper names (e.g. “Russian”).

The readings distinguished for the base class LOCATION are the literal reading, general metonymic patterns (Object-for-Name, Object-for-Representation), LOCATION-specific patterns (Place-for-People, Place-for-Event, Place-for-Product) as well as mixed readings and unconventional metonymies under the tag “othermet”. Place-for-people metonymies have the subtypes CapGov, OFF, ORG, POP and PER. Capital cities have the additional possible metonymic reading Capital-for-Country (see Appendix B). The readings are shown in Figure 3.1.

3.1.1 General description and literal reading

Many locations (e.g. countries and cities) describe a territory as well a political and economic or administrative unit inclusive of a government, inhabitants, culture and organisations. These readings are *literal* readings.

Test. The replacement test is to replace x by the following:

- (in) the geographical region (territory) known as x / (in) the geographical region (territory) occupied by x / the state/district/city x

As the replacement test here is sometimes non-conclusive, we give also grouped examples of literal readings as follows.

3.1.1.1 Geographical Descriptions

- **Description of the biological/geographical/geological conditions/phenomena of a place**, e.g. “*Germany* has lots of mountains.” or “*Malta* is naturally rocky, windy, tree-less.”
- **Describing the geographical location of a place.** “*Malta* lies 220 miles north of Libya.” or “*Germany* is east of France.”

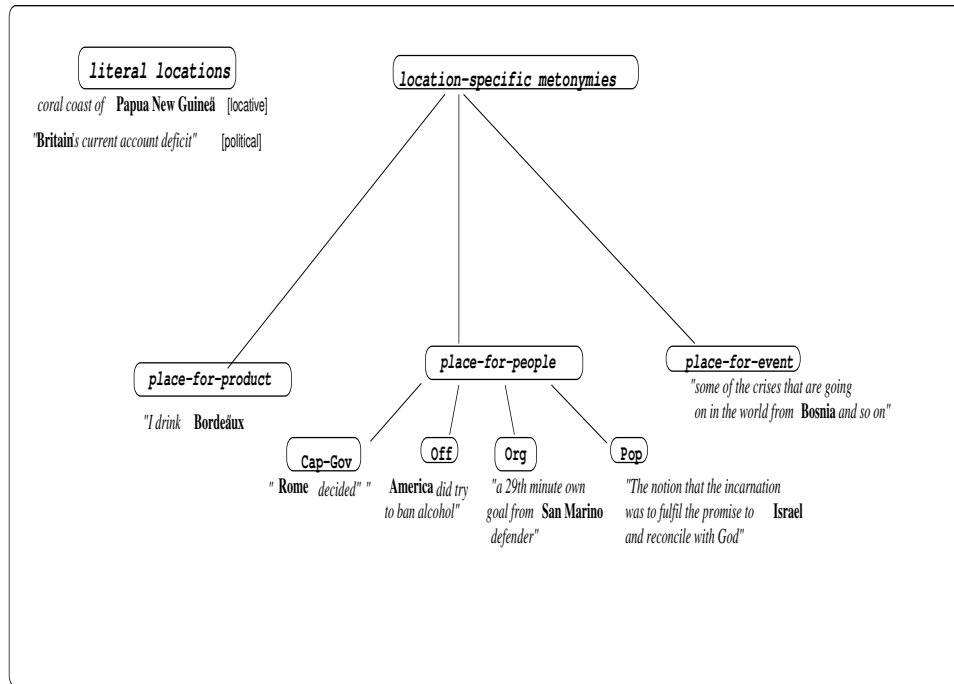


Figure 3.1: LOCATION specific metonymic patterns

3.1.1.2 Associations to natural objects, places, people and organisations

Associations are normally expressed via the following constructions, although other possibilities may be found: genitives with the location as the modifier (*“Germany’s government”*), “of”-constructions (*“the government of Germany”*), adjective constructions (*“German government”*), prepositional constructions (*“the government in Germany”*), appositives and associations via parenthesis (*“Bonn, Germany”* or *“Bonn (Germany)”*) and “have” constructions (*“Germany has a democratically elected government”*).

- **Associations between the place and natural objects or non-human animals and plants.** Examples are *“Italy’s trees”*, *“French coast”* or *“German coal”*.
- **Association to an embedded place.** *“the Spitfire Squadron at Luqa, Malta”* or *“Scotland’s capital”*, *“Italy’s south”*. The **embedding region** (here Malta, Scotland, Italy) is to be annotated as literal.

The annotation of the embedded region depends on the predicates it modifies: so in *“Scotland’s capital sleeps”* “Scotland” is to be annotated as literal whereas “capital” is to be annotated as a Place-for-People metonymy.

- **Associations between people and the place.** People include citizens, representatives, delegates, residents etc. The associations include location, representing, coming from etc. Examples are *“the president of the US”*, *“citizens/residents of Germany”*, *“the delegates of Germany at the UN”*, *“a French lawyer/Nobel prize winner”*, *“visitors of France”*, *“Japan’s homeless”*, *“Scotland has many homeless.”* or *“Director Hans Blix (Sweden)”*. It also includes associations between states and athletes *“... seven shots behind the leader Frederick Larsson of Sweden”*. See also Appendix B for special rules for the names of capital cities.

- **Associations between organisations and the place.** Organisations include companies, government organisations etc. (for a full list of what we consider as organisations see also Appendix A). Examples are “*the government of the US*”. “*Sweden’s fifth largest conglomerate*”, or “*Sweden’s hospitals*”. It includes the associations between states and sports teams. See also Appendix B for special rules for the names of capital cities.

Associations to artifacts, abstract and cultural attributes as well as economic attributes are dealt with in Section 3.1.3.

3.1.1.3 Other examples

- **Political, economic or diplomatic relationships between two locations.** Example: “*relationships/conflicts between Scotland and England*”.¹
- **Memberships in political or international organisation; signatories; observers.** “*Germany is a member of the EC*” as well as “*Germany joined/left the EU*”.² This also includes two stock phrases “to be a signatory of” and “to be an observer of” (e.g. “*Malta will be an observer of the peace process*”). NOTE however that “*Germany signed*” or “*Germany observes*” or “*Observers: Germany, Sweden...*” would be Place-for-People readings.
- **Describing the place where an action/event takes place or which is the physical destination or source of an action:** “*To write in Great Britain*”, “*executions in Iran*”, “*concert/summit in England*”, “*to live in UK*”. It also includes physical movement towards a place or physical actions directed towards a place, e.g. “*to smuggle into Denmark*” or “*a visit/move to France*”.
- **Target of military or political actions that are directed at the state as a whole..** “*To boycott Iraq*”, “*sanctions against Iraq*”, “*defence of Kuwait*”, “*liberate/annex Kuwait*” or “*to divide Germany*” as long as the context does not indicate that only one specific aspect of the country is the object of the action (e.g. liberating only the women of Kuwait). Most of the time these actions are at least partly also directed towards the **territory**.
- **Emotional attitudes towards a place and judgements of a place as a whole.** “*I love Italy*” or “*France is a great country.*” However, if the judgement is only of a particular part of the place or an action connected with the place, then a metonymic reading has to be considered. So, e.g. if “*I love Italy*” means *I love travelling to Italy* then a Place-for-Event metonymy is to be annotated.
- **Treating the country as an amalgam of all its components.** This is rare and results in general readings like “*Haiti is not America.*”. Another example would be “*Afghanistan’s problems*” if the problems encompass a wide range of problems affecting the whole of Afghanistan. Another Example: “*resolutions on Kuwait*”.

Remarks, Hints and possible confusions. Sometimes it is hard to distinguish literal readings from Place-for-People readings. See Section 3.1.3 for the following

¹Note that in certain contexts the relationship might be restricted to a particular group, e.g., to the two football teams of England and Scotland. This is then a metonymic reading.

²Although this looks like an action and therefore like a Place-for-People reading, the replacement “the government of Germany joined” is incorrect. Therefore we assign a literal reading.

cases: Associations between place and economic attributes, situations and actions; associations between artifacts/attributes. In addition, names of capital cities have their own problems, see Appendix B.

Water bodies include islands or bases etc. Thus, e.g. *“He was stationed in the Pacific”* will be annotated as literal.

Marking. Mark the value “literal” for the attribute “reading”.

3.1.2 Place-for-(set of)-People.

A location expression can stand for people/groups/organisations associated with it. Whenever this is the case, assign the tag **place-for-people** to the noun. Often used relationships between places and people are: people inhabit the place normally or habitually, people represent it, people/organisations are legally affiliated to it, e.g. by being citizens of the entity, by an organisation being founded in it etc. The relations can also involve several steps of connection.

We will treat organisations (e.g. companies or institutions) as sentient agents and not distinguish between **representatives** of the organisation/group and the organisation itself. Indeed, in many cases it is almost impossible to tell the difference and only relevant for a very deep understanding of the text.

Test. The replacement test is: replace x by

- people/group/organisation associated with/inhabiting/frequenting/representing
x

Examples.

1. **Instigators of actions that involve physical actions.** (e.g. speaking, moving muscles, applauding, signing etc.) E.g. *“Germany signed...”*, *“Iraq invaded”*, *“The whole city is in the streets”*. This includes the instigators of **military** actions (e.g. *“Russian intervention”*) Although we treat the target of a political or military action as a literal reading, we treat the instigator as metonymic, e.g. in *“Iraq invaded Kuwait”*, “Kuwait” would be literal whereas “Iraq” would be a Place-for-People metonymy.
2. **The instigators of political decisions, attitudes and actions. The makers of political communications.** E.g. *“Malta endorsed a series of proposals ...”*, *“a directive from Brussels”* as a directive from the EU which has the seat in Brussels. *“Cyprus and Malta ... submitted formal applications for membership of the EC.”* (although only states are members of an alliance, the government or other representatives of the state submit an application). *“Lesotho, Madagascar and Mauritius are in the process of implementing schemes...”*.
3. **The experiencers of emotional or mental states.** This includes emotional states like fear, panic, joy or anxiety. Example: *“Washington is insensitive to the needs of people.”*
4. **The recipient/beneficiary of actions or target of programmes that are not military or political actions directed at the state as a whole and that can only be directed at people** e.g. *“Aids programme for Uganda”* or *“to play for Germany”* (meaning to play for the Germany football team), *“to accuse Nigeria”* (for accusing the government or military). Note that most political and military actions are seen as targeted at

the state as a whole (especially if the territory of the state is targeted as well as in “annex Kuwait”) and are therefore annotated as literal.

We will also distinguish between subtypes depending on the kind of people meant. The subtype has to be annotated as a value for the attribute “subtype”. See Appendix A for the subtype specifications.

Marking. Mark the value “metonymic” for the attribute “reading”, the value `Place-for-People` for the attribute “metotype” and, if possible, the subtype as the value for the attribute “subtype”.

3.1.3 Remarks and possible Confusions.

We describe possible confusions between `Place-for-People` metonymies and literal readings.

3.1.3.1 Stock Phrases

Note that in some cases the subject of a verb denoting an action or a mental/emotional state that normally calls for a `Place-for-People` reading, can be used as a stock phrase. An example is “*Sweden has recorded a 10 percent fall in inflation.*”. Here “has recorded” is not used as an active verb anymore, but as a stock phrase, meaning “Sweden has a 10 percent fall in inflation”. We annotate these readings then as literal. To decide on what is a stock phrase we can often use dictionaries: so the PONS dictionary gives “record a measurement” as an idiomatic phrase.

Another example is a metaphoric phrase, expressing e.g. a long-term obligation. An example is “*Sweden is committed to phasing out nuclear power by 2010.*”. This expresses long-term obligation in contrast to the attitude of particular people or a particular government and should be annotated as literal.

3.1.3.2 Association between artifacts/attributes and a place.

Here the following guidelines hold:

1. If the attribute or artifact is an official attribute of a location, then a literal reading is assigned. Examples are: “*The language of **Germany***” or “*the religion of **Afghanistan***”. It also includes e.g. flags, laws as in “***Sweden**’s laws/flag/constitution/social system*”.
2. If the attribute/artifact is normally associated with people and not an official attribute of a location, then use a `Place-for-People` reading. This is e.g. true for languages and religion if the place does not have an official language or religions (e.g. “***Italy** is catholic*”). Other examples are local costumes like in “*costumes of **Peru***” or traditions /rites. If a verb denoting a physical activity is involved, e.g. “***Peru** wears black*” or “***Peru** speaks ...*” we treat it as a `Place-for-People` metonymy in any case.
3. If the artifact is just an artifact that is situated in the place or comes from the place, then use a literal reading. Examples are “*a building/canal in **France***” or “***German** films/cars/books*”.

3.1.3.3 Economic situations, attributes and actions

The following guidelines hold:

1. **Economic situations and attributes of a location as a whole** are annotated as literal readings. This holds e.g. for the national income, assigning expenditure etc. Examples are “*Afghanistan has high debts*” or “*high unemployment in Germany*” or “*Britain’s economy*”.
2. **Economic actions can be targeted** at the place as a whole, e.g. “*exports to Kuwait*” or “*boycott Iraq*”. In these cases they cannot be replaced by a people reading (thus, e.g. “exports to organisations/people” is not correct) and are thus easy to recognise. The beneficiaries/targets of these actions are annotated as literal. Some actions, though, can only be targeted at people, e.g. buying or selling. Then we treat the reading as a Place-for-People reading.
3. **Instigators of economic actions** are normally people/business organisations and we have a Place-for-People reading. Examples are “*a joint project between Bolivia and Argentina*” or “*Kuwait sells oil*”.

3.1.4 Place-for-Event

Sometimes places are strongly associated with specific events that happened in the place. Typical events are accidents, catastrophes (e.g. avalanches, fires, explosions...), military events (wars and battles, invasions, war crimes, terrorist attacks), political events and meetings (e.g. summits), scandals, social events (e.g. weddings, funerals, parties, ceremonies), sports events (e.g. Olympic Games, races) and cultural events (e.g. concerts, exhibitions, festivals).

Apart from specific events we also include the general situation in the place as well as actions connected with the place.

Test. The replacement test is: replace x by one of the following

- event that takes/took place in x
- situation in x
- action in/directed towards x

Examples.

- “*He is still shocked by Vietnam*”. (for the war in Vietnam)
- “*Vietnam veteran*”
- “*His four case studies, of Mauritius*” (context-dependent; might be situation in Mauritius)
- “*Watergate changed our politics.*”
- “*He won Wimbledon*”.
- “*Sweden on two herrings a day*: a book title meaning “Travelling through Sweden on ...”

Marking. Mark the value “metonymic” for the attribute “reading”, the value Place-for-Event for the attribute “metotype”. In addition, make a note on the type of event in the attribute “notes”.

Remarks and possible Confusions. In some cases the literal reading might be true, but an event reading might also be true. We treat the event reading as an

inference from the literal reading and world knowledge and annotate the reading itself as literal. An example will make this clear: “*He wanted to get his family out of **Bosnia***”. Here two readings are true: so, he really wanted to get his family away from Bosnian territory and, by doing so, he wants to get his family away from the war there. We give precedence to the literal reading.

3.1.5 Place-for-Product

Very rarely a place name is used for products that have been produced in the place. This scheme is restricted to artifacts. These artifacts are supposed to be sold in shops or be exported. In cases of doubt, see whether the intended reading is a “product” or “artifact” in WordNet.

Test. Replace x by

- the product(s) y that was (were) produced in x.

Examples.

- “*I bought a real **Meissen***.” (City for the porcelain made there.)

Hints and Possible Confusion. This pattern does not occur with common nouns.

Marking. Mark the value “metonymic” for the attribute “reading” and the value “Place-for-Product” for the attribute “metotype”. In addition try to identify the kind of product meant in the attribute “notes”.

3.1.6 Unconventional metonymies

In principle, places can stand for everything associated with them. An example is “*The thing about the record is the influences of the music. The bottom end is very New York/**New Jersey** and the top is very melodic*”. Here “New Jersey” stands for the typical tunes from New Jersey.

Marking. Mark the value “metonymic” for the attribute “reading” and the value “othermet” for the attribute “metotype”.

3.2 Annotation scheme for the semantic class ORGANISATION

At the moment, the annotation scheme for ORGANISATION covers the proper names of business organisations and companies. (We follow the MUC-7 annotation guidelines for determining the extent of an organisation name. It has not been determined yet whether it is applicable for appropriate common nouns like “firm” or “company” as well. It has also not been determined whether it is applicable for organisation proper names, which are not business organisations. We assume that most points are valid for these words as well, though. In this annotation scheme, examples taken from the BNC are unmarked, whereas examples that were made up are marked with an asterisk.

The readings distinguished for the base class ORGANISATION are the literal reading, general metonymic patterns (Object-for-Name, Object-for-Representation), ORGANISATION-specific patterns (Organisation-for-Members, Organisation-for-Facility, Organisation-for-Product, Organisation-for-Event, Organisation-for-Index) as well as mixed readings and unconventional metonymies under the tag “othermet”. Organisation-for-Members metonymies have the subtypes OFF, POP and OTHMEMBER (see Appendix C)

3.2.1 General description and literal reading

The *literal* reading of organisations describes references to the organisation in general, where organisation is seen as a structure consisting of organisation members that do speak with a collective voice and has a charter, statute or defined aims.

Test. The replacement test is to replace x by the following:

- (in) the organisation (company/firm) known as x

Examples of literal readings are:

3.2.1.1 Associations between organisations and its members, other organisations, products and services, headquarters/seats

- **Description of the make-up/organisation of an organisation**, e.g. “**BMW employs 3000 workers.*” or “*The Scientific Affairs Board of the BPS ...*”. This includes associations between an organisation and one of its members or substructures, e.g. “*Mr. Nigel Whittaker of Kingfisher*”. Such associations can be expressed via of-genitives, Saxon genitives, compounds (“{*IBM executive*” as well as appositives and prepositional constructions or via verb phrases, e.g. “**I work for IBM*”. It also includes relations between a suborganisation and an organisation e.g. “*the heavy engineering subsidiary of ENI*”. It also includes relations between delegates that represent the whole organisation and the organisation “**delegates of IBM*”.
- **Description of headquarters/seats or foundation of the organisation**. “**The Red Cross was founded in 1789*” or “**The headquarters of IBM are in Boston.*”.
- **Associations between organisations and their products or services**. This includes associations between organisations and the products they produce. Examples are “*Intel’s Indeo video compression hardware*”, “*IBM workstations*”. This also includes associations between an organisation and its market value/index, e.g. “**IBM shares*”.

- **Relations between different organisations or organisations and GPEs.** Official relations between different organisations are also to be annotated as literal. Examples are: “*Oregon-based **Mentor Graphics Inc.** has teamed with **Toshiba Corp.**” or “*the brand-new **Sun Microsystems Inc** and **Adobe Computer Systems Inc** alliance*”. This also includes sponsorships as in “*his first outing with his **Williams team***”. This also includes memberships of organisations in other organisations, e.g. “***IBM** joined the Association for Computer companies.*”. Please note that this is not true for memberships that only apply to non-representative/official parts of the organisation, e.g. workers of a company joining a union.*

3.2.1.2 Treating the organisation as an amalgam of all its components

This is rare and results in general readings like “***IBM** is not Microsoft.*”. Another example would be “***IBM’s** few current successes*”.

3.2.1.3 Organisation as target of actions or attitudes.

- **Target of economic or political actions that are directed at the organisation as a whole.** Examples are “***buy Nestle***”, “***sue Microsoft***”, “*case against **IBM***” as long as the context does not indicate that only one specific aspect of the organisation or a product of the organisation is the object of the action (e.g. only buying Nestle products). Another example is “*the resources to challenge **Microsoft***”

In case of boycotting, for example “*to boycott **Nestle***”, an organisation-for-product reading is indicated, as the sentence is equivalent to “*to boycott **Nestle goods***”.

- **Emotional attitudes towards an organisation and judgements of an organisation as a whole.** “***I hate Microsoft***” or “***Microsoft** is a great company.*” However, if the judgement is only of a particular part/product of the organisation or an action connected with the organisation, then a metonymic reading has to be considered. So, e.g. if “***I hate Microsoft***” means *I hate operating systems produced by Microsoft* then an Organisation-for-Product metonymy is to be annotated.

3.2.1.4 Commercial Actions

This applies to the agents of standard commercial actions typical of companies, especially for a company towards a product. For example, “*a hand-held CD-Rom player developed by **Apple** and **Toshiba***”, “*competing offerings from **Sun Microsystems Inc***”, “***Sprint** provided a managed network*”. Predicates possibly involved are “sell”, “buy”, “deliver”, “develop”, “produce”, “make”, “launch”. Actions directed not towards products but towards other companies/businesses can be expressed by predicates such as “buy”, “sell”, “spend”, “invest”. These are also to be annotated as “literal” in most cases. Examples: “***Intel** will have spent \$5,000m on it.*”, “***Sun** acquired that part of **Eastman-Kodak Co’s** Unix subsidiary*”.

Marking. Mark this reading as “literal” for the attribute “reading”.

3.2.2 Organisation-for-Members

An organisation name can stand for members associated with it. Whenever this is the case, assign the tag `organisation-for-members` to the noun. Often used relationships between organisations and members are: people who work for the

organisation or people who represent the organisation as well as suborganisations of the organisation, e.g. member states of an international organisation.

Test. The replacement test is: replace x by

- people working for and/or representing x; OR member states of x; OR member organisations of x

Examples.

1. **Instigators of communicative actions.** Examples are “speaking”, “announcing”, “accusing”, “saying”, “declaring”, “discussing”, “advertising”. Included are also legal communications/actions such as “signing”, “writing”, “sueing”, “bidding”, “fighting”, “agreeing”. Examples: “*discussions with Intel*”, “*IBM [...] recently announced*”, “*Sun said*”, “*signing connection agreements with Enron*”. “**IBM and Microsoft ... submitted formal applications for membership of the Association for .*” (although the whole organisations are members of an alliance, representatives of the organisation submit an application).
2. **The instigators of decisions, attitudes and plans/goals.** Examples are “*IBM’s own development goals*”, “*Nestle had to abandon that*”. Predicates usually involved are “decide”, “want”, “think”, “have to”.
3. **The experiencers of emotional or mental states.** This includes emotional states like fear, panic, joy or anxiety. Example: “*BT can be pleased*”
4. **The recipient/beneficiary of actions or target of programmes that are not political or economic actions directed at the organisation as a whole and that can only be directed at people** e.g. “*to play for IBM*” (meaning to play for the IBM football team), “*to accuse IBM*” (for accusing the officials).

We will also distinguish between subtypes depending on the kind of members meant. The subtype has to be annotated in the memberssubtype field. See Appendix (C) for the subtype specifications.

Marking. Mark the value “metonymic” for the attribute “reading” and the value `organisation-for-members` for the attribute “metotype”. If possible, identify the subtype as the value for the attribute “subtype”.

3.2.3 Remarks and possible Confusions.

We here go into the details of possible confusions between `Organisation-for-Members` metonymies and literal readings.

3.2.3.1 Stock Phrases, metaphors and official actions of the organisation

Note that in some cases the subject of a verb denoting an action or a mental/emotional state that normally calls for a `Organisation-for-Members` reading, can be used as a stock phrase. An example is “**IBM has recorded a 10 percent fall in income*”. Here “has recorded” is not used as an active verb anymore, but as a stock phrase, meaning “IBM had a 10 percent fall in income”. We annotate these readings then as literal. To decide on what is a stock phrase we can often use dictionaries: so the PONS dictionary gives “record a measurement” as a special phrase.

Another example is a metaphoric phrase, expressing e.g. a long-term obligation. This then calls for a literal reading of the subject (organisation). An example is “**IBM is committed to x*”. This expresses long-term obligation in contrast to the attitude of particular people or particular officials. Another example: “*Computer Science Corp has won a multi-year facilities management contract*”

3.2.3.2 Association between artifacts/attributes and an organisation

Here the following guidelines hold:

1. If the attribute or artifact is an official or global attribute of the organisation, then a literal reading is assigned. Examples are “**The flag of the Red Cross*”. It also includes e.g. rules as in “**IBM’s charters/pension system.*” or official languages, e.g. of the EU, as well as official uniforms e.g. of McDonalds. This category also includes statutes/charters that identify official aims, e.g. “**The statutes of the UN*”. As global features, an example is “*the size of AT&T and IBM*”.
2. If the attribute is ascribed to the organisation as a unique legal entity, then the reading is a literal reading. An example is “**Microsoft is powerful*”. The test is whether these attributes apply to the organisation even if the members of the organisation does not share the attributes. Thus again a replacement “the/all/most members of Microsoft are powerful” distorts the truth semantics of the sentence.
3. If the attribute/artifact is normally associated with people and not an official attribute of an organisation, then use a **Organisation-for-Members** reading. This is e.g. true for non-official languages/religions/costumes/customs/attitudes. If a verb denoting a physical activity is involved, e.g. “**NATO wears black*” or “**IBM speaks ...*” we treat it as an **Organisation-for-Members** metonymy in any case.
4. If the artifact is just an artifact that comes from the organisation, belongs to the whole organisation or is a product of the organisation, then use a literal reading. Examples are “*IBM workstations*”, “**Red Cross building*”, or “**Red Cross van*”.

3.2.3.3 Economic situations, attributes and actions

The following guidelines hold:

1. **Economic situations and attributes of an organisation as a whole** are annotated as literal readings. This holds e.g. for the organisation’s income, expenditure etc. Examples are “**IBM has high debts*” or “**low salaries at IBM*” or “*Toyota’s financial profit*” or general descriptions of economic situations “**IBM is rich*”.³
2. **Economic actions can be targeted** at the place as a whole, e.g. “*buy IBM*” (see also Section 3.2.1.3). In these cases they cannot be replaced by a people/members reading and are thus easy to recognise. The beneficiaries/targets of these actions are annotated as literal. Some actions, though, can only be targeted at people. Then we treat the reading as a **Organisation-for-Members** reading.

³In some rare cases the attribute is clearly only an attribute of people (this is normally context-dependent and the people are *explicitly* mentioned before). We then assign a **Organisation-for-Members** reading.

3. **Instigators of economic actions:** see Section 3.2.1.4. For organisation-for-members, an example is “*Hitachi Ltd is to plug SoftBench*”.

3.2.4 Organisation-for-Facility

Frequently an organisation name is used for the facility in which it is housed. Facilities are in this case normally buildings.

Test. Replace x by

- the building(s) y that houses x.

Examples. The examples normally involve locations of persons or artefacts with relation to the organisation (e.g. being in front of, being at a street corner, being in the org/facility), or the part-of relation between artefacts and the org/facility.

- “**I entered **McDonald’s**.*”
- “**I entered **the McDonald’s**, where I normally eat breakfast.*”
- “**I stand in front of **IBM***”
- “****IBM** has new windows*”
- “**The protesters near the gate of **Microsoft***”
- “*the famed conversation in **Sainsbury’s***”

Marking. Mark the value “metonymic” for the attribute “reading” and the value organisation-for-facility for the attribute “metotype”.

3.2.5 Organisation-for-Product

Frequently an organisation name is used for products that have been produced by the organisation. This scheme is restricted to artifacts. These artifacts are supposed to be sold in shops or be exported. In cases of doubt, see whether the intended reading is a “product” or “artifact” in WordNet.

Test. Replace x by

- the product(s) y that was (were) produced by x.

Examples.

- “*Access from Windows and MS-DOS machines will be possible from **Oracle**.*”
- “*It was the largest **Fiat** anyone had ever seen.*”
- “***IBM** users*”
- “*boycott **Nestle***”

Marking. Mark the value “metonymic” for the attribute “reading” and the value organisation-for-product for the attribute “metotype”. In addition, try to identify the type of product in the attribute “notes.”

3.2.6 Organisation-for-Index

Frequently an organisation name is used for an index that indicates the value of the organisation. Examples of indices are shares or the stock index.

Test. Replace x by

- the index that indicates the value of x

Examples.

- “**IBM fell by three points.*”

Marking. Mark the value “metonymic” for the attribute “reading” and the value organisation-for-index for the attribute “metotype”.

3.2.7 Organisation-for-Event

Rarely organisations are strongly associated with specific events that happened at or to the organisation. Typical events are accidents, catastrophes (e.g. avalanches, fires, explosions...), military events (terrorist attacks), economic or political events and meetings (e.g. summits, business debates), scandals, social events (e.g. weddings, funerals, parties, ceremonies), sports events (e.g. Olympic Games, races) and cultural events (e.g. concerts, exhibitions, festivals).

Apart from specific events we also include the general situation as well as actions connected with the organisation.

Test. The replacement test is: replace x by one of the following

- event that takes/took place at x
- situation at x
- actions in/directed towards x
- event that involved x

Examples.

- “*In its **Philip Morris** decision in November 1987, the Court held...*”.
- “**He is still shocked by **Enron***”. (for the scandal involving Enron)
- “**His four case studies, of **IBM***” (context-dependent; might be situation at IBM)
- “****Enron** changed our opinion.*”

Marking. Mark the value “metonymic” for the attribute “reading” and the value organisation-for-event for the attribute “metotype”. In addition, make a note of the type of event in the attribute “notes”,

3.2.8 Unconventional metonymies

In principle, organisations can stand for everything associated with them apart from the above-mentioned (people/members, events, products, facilities, indexes). An example is “*funds [...] had been paid into **Barclays Bank***”. Please annotate such metonymies as “meto” in the field “reading” and “othermet” in the field metotype.

Marking. Mark the value “metonymic” for the attribute “reading” and the value othermet for the attribute “metotype”.

A Appendix: Subtypes of people readings

We specify between the subtypes `OFF`, `ORG`, `POP` and `PER` for the class `LOCATION`. For capitals of countries or states we have an additional subtype `CapGov`, where the capital stands for the government of the whole state.

If you can identify the subtype, specify it as the value of the attribute “subtype”, otherwise (or if several subtypes are included) leave it not-annotated.

Representatives. We regard representatives of an `ORG` or an official organisation (`OFF`) as an `ORG` or `OFF` reading according to the organisation. Thus, if a representative of a firm makes a decision for the firm, the subtype is `ORG`. If a representative of the government makes a decision for the government, it is `OFF`.

Groups of groups. We also include sets of subtypes (e.g. several organisations) in the original subtype. Therefore, if several firms act together `ORG` is still the type to be assigned. If several governmental organisations act together the type is still `OFF`.

Subgroups of groups. We also regard subgroups of groups/organisations as being of the original type. Therefore, if a division of a firm does something it is `ORG` as well. The same holds for the other types (e.g. divisions of governments and a battalion of the army are `OFF` as well).

A.1 CapGov

NOTE. This subtype exists only for capitals of countries or states!

There are cases where the capital of a country stands for the government of the whole country. A typical example is “*Washington opts out of Kyoto agreement*”, talking about the United States government.

This type can sometimes be confused with the `OFF` subtype. The `OFF` subtype for a city (e.g. Washington) is to be assigned if the administration/officials of the City are referred to (e.g., in a scenario where the city administration is building more streets in the city: “*Washington has decided to build more streets.*”)

A.2 OFF

This subtype comprises any officials (as individuals or as a group) **representing** a location as a political entity. This covers the government, including the executive: civil service, administration, and any representatives of these as well as monarchs, dictators, presidents, government ministries and ministers, and government departments (e.g. defence department, foreign office). We also include the army (as long as government controlled) as well as the police force, the secret service and other government agencies. We also include the legislative (parliament or senate) and the judiciary (courts, and judges).

For territories other than countries the corresponding officials are the administration of the place, e.g. city officials, local police force, the city chamber etc.

A typical example is “*Malta submitted a formal application for EC membership.*”

Note this type as the value `OFF` for the attribute “subtype”.

A.3 ORG

Sometimes the location is used to refer to a specific group or organisation originating from or residing in the place. A list of possible organisations or groups is given

below.⁴ Note that subgroups of groups (e.g. a battalion of a military force or a department of the university) count as groups as well. We also include representatives of these groups as well as sets of groups (e.g. several firms).

General

association
professional association
cooperative
collective (e.g. kibbutz)
club, society, guild, gild, lodge, order (e.g. bridge clubs, or also secret societies, e.g., the Klu Klux Klan)

Artistic

company, troupe (e.g. opera company, ballet company, chorus etc.)
musical organisation, musical group (e.g. duo, band, orchestra)

Sports

league (e.g. baseball league)
teams, squad (e.g. baseball team, crew)
club, society (e.g. tennis clubs)

Religious/Social

charity
religion, faith (e.g. catholic church)
denomination (e.g. Protestant denomination)
mission, missionary post, missionary station, foreign mission

Violent

gang, pack, ring, mob (e.g. an inner-city gang)
military or revolutionary groups (as long as not official, e.g. IRA)

Institutions

medical institution (e.g. hospitals)
educational institution (e.g. schools, universities)
academy, honorary society

Business/Finance

financial institution, financial organisation (e.g. insurance companies, investment companies or banks)
issuer
companies, business, concerns, corporations, firms (e.g. accounting firm or car dealers), manufacturers (e.g. shoe manufacturer)
brokerage
chain (of shops)
industry sector (e.g. fashion industry, or chemical industry, or entertainment industry)
consortium, pool, syndicate

Public

public service (e.g. telephone or gas companies)
carrier, common carrier (e.g. airline or railways)

⁴This list of organisations and institutions is mostly identical with the hyponyms (subconcepts) of “organisation” (Sense 2) in WordNet 1.6. Deleted are organisations which we classified as being official. In addition, some subconcepts were deleted, so as not to make the list too long.

fire brigade, fire company

Political

union, labour union, trade union, trades union, brotherhood

party, political party

alliance, coalition, alignment, alinement

world organisation, international organisation (e.g. EU, UN)

federation

committee, citizens committee

deputation, commission, delegation, delegacy, mission (e.g. diplomatic mission)

Others

agency (e.g. travel agency or detective agency)

pool – (an organisation of people or resources that can be shared; “a car pool”)

staff (of any of the mentioned groups or institutions)

jury

Please note that arbitrary assemblies of people at a given time as well as families are **not** treated as ORG. Examples are:

- “*A new directive from **Brussels***” meaning a new directive of the EC. (Note that this is context-dependent: the subtypes OFF (the administration of the city Brussels), CapGov (the administration of the country Belgium) and ORG (for EC) are all possible.
- (in a text about the Eurovision) “***Malta** gave 12 points to Ireland.*” (Malta for “jury of Malta”)
- “*Last week’s win over **Malta***”. (in a sports report)
- “***Malta** exported*”. Several business organisations.
- “***Paris** is introducing shorter skirts this season.*” (Fashion industry)

A.4 POP

Sometimes a place is used to refer to the majority of the people living in or frequenting the place. As a test you can replace x by “the (majority of) the population/people/electorate of x”. Typical examples are:

- “***Malta** fears*” Only people have emotions; however in some contexts this example might be OFF.
- “***Indonesia** voted its new government.*”

A.5 PER

Sometimes none of the above subtypes applies, although one can clearly identify the person or sets of persons referred to. Thus, in a scenario where a play is shown in a theatre in Nigeria and Germany, one might say “***Nigeria** liked it*”, meaning the audience in the theatre in Nigeria. This is not OFF and POP and also not a group/org as an audience is not an organised group. We annotate these readings as PER, referring to a family, an individual or an arbitrary assembly of persons.

B Appendix: Special rules for capitals

Capitals behave a bit differently than other location names as they have the additional possibility to stand for their country or the government of the whole country as well as the city and the city administration.

Please note these additional guidelines:

- If a people reading is clearly called for (normally if the capital name is the subject or direct-object of an action) , assign a `place-for-people` reading as the main pattern. Then decide on the basis of the subtype guidelines which people are meant. An example is “**Washington** decided...”. The context will then tell you the subtype `CapGov` or `OFF`.
- Otherwise, replace the capital name by the name of the corresponding country. If this gives the same meaning, think what reading that would be for the replaced country name and annotate correspondingly, using the `Capital-for-Country` metonymy. Example: “**Havana**’s foreign minister”. Here the reading is the same as “**Cuba**’s foreign minister. As “Cuba” would in such an example be annotated as literal, “Havana” must be annotated as a `Capital-for-Country` reading.
- **Compounds** are especially hard. Thus in a compound like “**Washington** officials”, the literal reading (“officials in Washington”) as well as the `Capital-for-Country` (officials representing the US) reading both make perfect sense in most contexts. If the context does not exclude the literal reading, we choose the literal reading.
- A similar conflict holds for readings like “The proposal was rejected in **Washington**”, where both the `CapGov` as well as the `literal` reading are true. We go for the literal reading if the context does not indicate otherwise.

C Appendix: Subtypes of members readings

We specify between the subtypes OFF, POP and OTHMEMBER for `organisation-for-members` readings.

If you can tell the subtype, specify it in the `memberssubtype` field, otherwise or if several subtypes are included leave it not-annotated. A common case is OFF and POP (e.g. the officials of an organisation and all members) being both included.

Groups of groups. We also include sets of subtypes (e.g. several suborganisations or member states) in the original subtype. Therefore, if several member states of the UN act together OTHMEMBER is still the type to be assigned. If several representatives of the organisation act together the type is still OFF.

C.1 OFF

This subtype comprises any officials (as individuals or as a group) **representing** an organisation as an official entity. This covers individuals e.g., managers, presidents, spokesmen as well as board members or committees that represent the whole organisation.

A typical example is “***IBM** submitted a formal application for membership in the Computer Science Association.*”

Note this type as OFF in the `peoplesubtype` field.

C.2 POP

This subtype comprises all/most of or all/most of the non-official, non-decision making members of an organisation. It only applies to people. A typical example is “***IBM** went on strike*”.

C.3 OTHMEMBER

If the organisation name is used to refer to any suborganisation or a member state only, please use OTHMEMBER. An example is “*I play for **IBM***” to refer to the football team of an IBM.