

Nominalizations and nationality expressions: A corpus analysis

Daniel Berndt¹ Gemma Boleda² Berit Gehrke³
Louise McNally³

University of Osnabrück

Universitat Politècnica de Catalunya

Universitat Pompeu Fabra

CLC 2009
July 21-23, 2009

A contrast and a question

Two strategies for expressing participant roles with event nominalizations:

- (1) **French** agreement to participate in the negotiations
- (2) agreement **by France** to participate in the negotiations

Question: What determines the choice between these options?

The goal of this talk: To present the initial results of a corpus study which will help answer this question.

Outline

- 1 Introduction
- 2 More on the data and issues
- 3 Previous proposals
- 4 Method
- 5 Analysis
 - Noun type
 - Concept stability

Delimiting the data

The adjective variant generally involves so-called **relational** adjectives (Bally, 1944; Levi, 1978):

(3) presidential, molecular, solar, cerebral, ...

Our study is limited to a subclass of these, the so-called **ethnic** adjectives, or **EAs** (Alexiadou and Stavrou, to appear):

(4) French, Spanish, American, South Korean, ...

Two basic uses for EAs

“Argument-saturating” use:

(5) **French** agreement to participate in the negotiations

“Classifying” use:

(6) **French** bread

See e.g. Bosque and Picallo (1996), though most theoretical work deals with the argument-saturating use only.

The theoretical issues these adjective raise

- We must account for the apparent synonymy between *French agreement* and *agreement by France* ...
 - ... but it would be desirable to give a single semantic analysis of the EA in its two basic uses.
- The adjective has been claimed to saturate an argument of the nominalization (Alexiadou and Stavrou, to appear) ...
 - ... but adjectives generally *modify*, they don't *saturate*.

An understanding of the distribution of the adjective and prepositional variants should shed light on these issues.

Previous proposals

- Ethnic adjectives are nominal in nature (Levi, 1978; Fábregas, 2007; Alexiadou and Stavrou, to appear).
- For illustration we will outline Alexiadou and Stavrou (A&S).

Details of A&S's analysis

EAs are morphologically deficient nouns that become adjectives in the course of the syntactic derivation.

- EAs originate in the specifier of the noun phrase they modify, where A&S assume the agent role is always assigned.
- Since EAs are not valued for case, a morphological requirement forces them to move to the specifier of a higher adjectival phrase; in this position they are spelled out as adjectives.

A&S's predictions

- The adjective must denote an agent (following Kayne, 1984, and others).
 - It is syntactically base-generated in the position where agents are base-generated.
- EAs are not gradable and cannot be coordinated with 'normal' adjectives, but only with other EAs.

(7) *French and strong agreement

- EAs are not proper adjectives but rather nouns underlyingly.
- EAs do not license anaphora (Postal, 1969).

(8) *The French_i agreement to ... They/It_i ...

- The underlying noun is morphologically deficient and is spelled out as an adjective.

Problems with A&S's analysis

- Our corpus study shows that EAs do not necessarily relate to agents.
 - Examples: *the Vietnamese arrivals*, *our French investment*, *the Indian debate*
- A&S themselves claim that the nominalizations combining with EAs in Greek lack argument structure (following Grimshaw, 1990).
 - Therefore, (1) cannot be a **real** argument saturating use.

Problems with A&S's analysis

- The coordination and gradability facts are not an argument against treating EAs as proper adjectives: there are other adjectives that do not coordinate freely, and many are not gradable (McNally and Boleda, 2004; Gehrke and McNally, to appear)
- Lack of anaphora to EAs is expected without assuming that they are underlying nouns.

Why a corpus study is useful

We provide basic data on the distribution of EAs vs. PPs:

- The sorts of nouns EAs vs. PPs occur with
- The variation in the head nouns that EAs occur with

This will ground our arguments against A&S's analysis and provide a departure point for a new theoretical analysis.

Data

- British National Corpus
- 49 nationality adjectives/nouns
 - list from Wikipedia
 - adjective (*French*) and proper noun (*France*) frequencies 1,000–30,000
- EA vs. PP examples, filtered as follows:
 - infrequent head nouns (≤ 24 occurrences)
 - nation-specific nouns (e.g., *reunification* for *German*), using entropy
 - for some purposes 45 manually selected nominalizations were used

Noun type

- Source for noun types: Top Concept Ontology (Álvez et al., 2008), WordNet based resource
 - Ontological category of head nouns.
 - Roughly equivalent to object-denoting (*wine*), event-denoting (*agreement*), and abstract entity-denoting (*idea*) nouns:
 - *French wine / wine from France*
 - *French agreement / agreement by France*
 - *French idea / idea in France*

Noun type

description	EA	PP	Total	% EA	%PP
object	19,816	10,866	30,682	64.6	35.4
event	8,813	10,898	19,711	44.7	55.3
abstract	2,168	1,986	4,154	52.2	47.8

- object-denoting nouns prefer the adjective construction
- event-denoting nouns prefer the prepositional construction
 - effect even stronger if only nominalizations are considered:

	EA	PP	Total	% EA	%PP
nominalizations	492	1,788	2,280	21.6	78.4

Noun type

description	EA	PP	Total	% EA	%PP
object	19,816	10,866	30,682	64.6	35.4
event	8,813	10,898	19,711	44.7	55.3
abstract	2,168	1,986	4,154	52.2	47.8

- object-denoting nouns prefer the adjective construction
- event-denoting nouns prefer the prepositional construction
 - effect even stronger if only nominalizations are considered:

	EA	PP	Total	% EA	%PP
nominalizations	492	1,788	2,280	21.6	78.4

Noun type

description	EA	PP	Total	% EA	%PP
object	19,816	10,866	30,682	64.6	35.4
event	8,813	10,898	19,711	44.7	55.3
abstract	2,168	1,986	4,154	52.2	47.8

- object-denoting nouns prefer the adjective construction
- event-denoting nouns prefer the prepositional construction
 - effect even stronger if only nominalizations are considered:

	EA	PP	Total	% EA	%PP
nominalizations	492	1,788	2,280	21.6	78.4

EAs are used for stable concepts

- The adjective construction is used for stable concepts.
- *A war has started in Spain ...*
 - *The war in Spain ...*
 - *The Spanish war ...*
- We should observe **less variety** in the head nouns for the adjective construction.

Nation	EA	PP	Union	% EA	%PP
France	15	30	34	44.1	88.2
Germany	17	26	32	53.1	81.3
Japan	12	23	28	42.9	82.1
...					

EAs are used for stable concepts

- The adjective construction is used for stable concepts.
- *A war has started in Spain ...*
 - *The war in Spain ...*
 - *The Spanish war ...*
- We should observe **less variety** in the head nouns for the adjective construction.

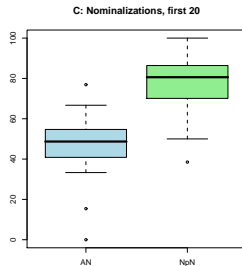
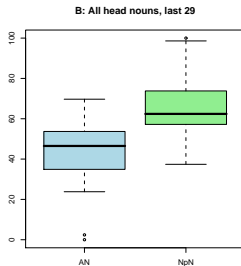
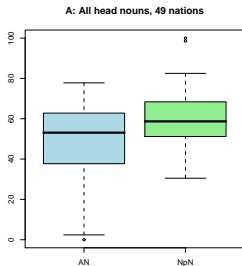
Nation	EA	PP	Union	% EA	%PP
France	15	30	34	44.1	88.2
Germany	17	26	32	53.1	81.3
Japan	12	23	28	42.9	82.1
...					

EAs are used for stable concepts

- The adjective construction is used for stable concepts.
- *A war has started in Spain ...*
 - *The war in Spain ...*
 - *The Spanish war ...*
- We should observe **less variety** in the head nouns for the adjective construction.

Nation	EA	PP	Union	% EA	%PP
France	15	30	34	44.1	88.2
Germany	17	26	32	53.1	81.3
Japan	12	23	28	42.9	82.1
...					

Adjective construction is used for stable concepts



Conclusions

- Concept stability is an important factor in the choice between EAs and PPs.
 - (Classificatory use a special case of concept stability?)
 - A priori, EAs+nominalizations are unlikely candidates for stable concepts
(cp. *French bread* vs. *French agreement*)
 - This explains why nominalizations strongly prefer PPs.
- If EAs were able to saturate arguments, we would expect them to appear more frequently with, and with a wider variety of, nominalizations.
 - Our data thus argue against analyses like A&S's.

Nominalizations and nationality expressions: A corpus analysis

Daniel Berndt¹ Gemma Boleda² Berit Gehrke³
Louise McNally³

University of Osnabrück

Universitat Politècnica de Catalunya

Universitat Pompeu Fabra

CLC 2009
July 21-23, 2009

- Alexiadou, A. and Stavrou, M. (to appear). 'Ethnic adjectives as pseudo-adjectives: A case study in syntax-morphology interaction and the structure of DP'. *Studia Linguistica*.
- Álvarez, J., Atserias, J., Carrera, J., Climent, S., Laparra, E., Oliver, A., and Rigau, G. (2008). 'Complete and consistent annotation of wordnet using the top concept ontology'. In *6th Language Resources and Evaluation Conference (LREC 2008)*. Marrakech, Morocco.
- Bally, C. (1944). *Linguistique générale et linguistique française*. Berne: A. Francke.
- Bosque, I. and Picallo, C. (1996). 'Postnominal adjectives in Spanish DPs'. *Journal of Linguistics*, 32:349–386.
- Fábregas, A. (2007). 'The internal syntactic structure of relational adjectives'. *Probus*, 19.1:135–170.
- Gehrke, B. and McNally, L. (to appear). 'Frequency adjectives and assertions about event types'. In *Proceedings of SALT 19*. Ithaca, NY: CLC Publications.

- Grimshaw, J. (1990). *Argument Structure*. Cambridge, MA: MIT Press.
- Kayne, R. (1984). *Connectedness and Binary Branching*. Dordrecht: Foris.
- Levi, J. (1978). *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.
- McNally, L. and Boleda, G. (2004). 'Relational adjectives as properties of kinds'. In O. Bonami and P. C. Hofherr (eds.), *Empirical Issues in Syntax and Semantics*, vol. 5. <http://www.cssp.cnrs.fr/eiss5>, 179–196.
- Postal, P. (1969). 'Anaphoric islands'. In *Proceedings of the Fifth Regional Meeting of the Chicago Linguistics Society*. 205–239.