

A Corpus of Indefinite Uses Annotated with Semantic Functions: Documentation

Ana Aguilar-Guevara¹, Maria Aloni², Tikitu de Jager², Angelika Port², Radek Šimík³, Stephanie Solt², and Machteld de Vos²

¹Utrecht University

²University of Amsterdam

³University of Potsdam

November 1, 2012

1 Introduction

The *Corpus of Indefinite Uses* is an output of the NWO *Vidi* project **Indefinites and beyond. Evolutionary pragmatics and typological semantics** (PI: Maria Aloni). It makes available to interested researchers data collected and annotated in the course of a cross-linguistic synchronic and diachronic corpus study of indefinite expressions in six different languages:

- German *irgend* series
- Italian *(uno) qualunque*
- Spanish *cualquiera*
- Dutch *wie dan ook*
- Czech *kterýkoli*
- English *any*

The main goal of the study was to verify the distribution of these indefinites synchronically and to attest to their historical development. The methodology used was a form of functional labeling which combines both context (syntax) and meaning (semantics), using as a starting point Haspelmath's (1997) functional map.

The present document serves as an introduction to the data set that formed the basis for this study. We outline here the theoretical background to the corpus study, the methodology for functional classification of corpus examples, and the structure and capabilities of the online database, and then summarize the key synchronic and diachronic results for the six items investigated.

2 Theoretical Background

It is well known that the use of expressions with existential meaning (e.g. plain indefinites like English *somebody*, German *jemand* and Czech *někdo*) can give rise to different pragmatic effects. Although the semantic representation of *somebody* in (1) and (2) is identical, (1) comes along with a **free choice implicature** (each individual is a permissible option) and (2) with an **ignorance implicature** (the speaker does not know who called):

- (1) You can invite somebody.
- (2) Somebody called.

From a typological perspective, many languages have developed specialized forms for such enriched meanings, such as **free choice indefinites**¹: Italian *-unque*-series, Czech *koli*-series, Dutch *dan ook*-series, Spanish *cualquier*-series, and **epistemic indefinites**²: Russian *to*-series, Czech *si*-series, German *irgend*-series, Spanish *algún*-series.

Following Grice’s seminal work, the main hypothesis that motivates the present research is that these different indefinite forms have emerged as the result of a process of conventionalization (or fossilization) of an originally pragmatic inference.

It may not be impossible for what starts life, so to speak, as a conversational implicature to become conventionalized. (Grice 1975:58)

In languages with Epistemic Indefinite (EI) forms, inference (3-c), pragmatic in origin, has been integrated into the semantic content of sentences like (4-a).

- (3) *Plain indefinite (German)*
- a. **Jemand** hat angerufen.
somebody has called
 - b. Conventional meaning: Someone called
 - c. Ignorance implicature: The speaker does not know who
- (4) *Epistemic indefinite pronoun (German ‘irgendjemand’)*
- a. **Irgendjemand** hat angerufen.
somebody:UNKNOWN has called
 - b. Conventional meaning: Someone called and the speaker does not know who

In languages with distinctive Free Choice (FC) forms, inference (5-c), pragmatic in origin, has been integrated into the semantic content of sentences like (6-a).

- (5) *Plain indefinite (Spanish)*
- a. Puedes traer **un** libro.
can:2SG bring:INF a book
 - b. Conventional meaning: You can bring a book
 - c. Free choice implicature: Each book is a possible option
- (6) *Free choice determiner (Spanish ‘cualquier’)*
- a. Puedes traer **cualquier** libro.
can:2SG bring:INF any book
 - b. Conventional meaning: You can bring a book and each book is a possible option

In this project, a number of cross-linguistic synchronic and diachronic studies have been combined in order to substantiate this hypothesis. The synchronic studies were intended to determine what has been fossilized, the diachronic studies how this has happened.

In the synchronic research we studied the following indefinite forms: German EI *irgend* series, Czech FC *kteřjkoli*, Italian FC (*uno*) *qualunque*, Spanish FC *cualquiera*, Dutch FC *wie dan ook* and English FC *any*. The main goal of this research was to understand which part of the meaning of the indefinite form is fossilized and to develop some hypotheses on how this might have happened diachronically. In the diachronic corpus research we studied the historical development of three of these indefinite forms: Spanish *cualquiera*, Dutch *wie dan ook* and German *irgend(ein)*.

Below we describe the methodology developed for these corpus studies, and the structure and functionality of the online database where the data are housed, and we summarize the key findings of the synchronic and diachronic portions of our research.

3 Corpus study: diagnostics and methodology

In the synchronic and diachronic studies we classified randomly selected occurrences of each indefinite according to a number of categories. The annotation was carried out by six annotators (one per language), all with linguistic training. Five of these annotators (for German, Italian, Spanish, Dutch and Czech) met regularly to compare their results and share their experience with the annotation instructions. The sixth annotator (for English) completed the

¹E.g. Dayal (1998), Giannakidou (2001), Menéndez-Benito (2010).

²E.g. Alonso-Ovalle and Menéndez-Benito (2010), Jayez and Tovena (2006), Kratzer and Shimoyama (2002).

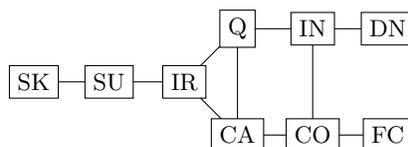
task at a later date. In addition to annotation by linguistically trained annotators, the English data were separately annotated by non-expert annotators (graduate students at the ILLC, University of Amsterdam), following a set of guidelines developed in the earlier stage of the project. Results of this work, including inter-annotator agreement scores, are reported in Aloni et al. (2012), and will not be discussed here.

The starting point for the identification of the relevant categories was Haspelmath’s (1997) implicational map.

3.1 Haspelmath semantic map

Haspelmath’s (1997) typological survey identified 9 main functions (context/meaning) for indefinite forms organized in an implicational map. Haspelmath proposes that an indefinite will always express a set of functions that are contiguous on the map (where two functions are contiguous iff they are connected by a line). One prediction is that items which acquire new functions will develop first those functions that are adjacent to the original function.

(7) *Haspelmath’s map*



(8) *Functions on the map*

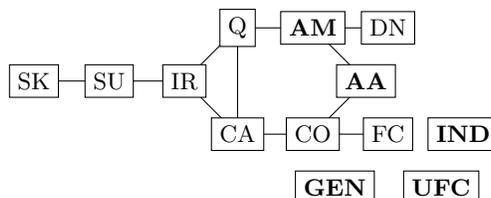
	Abbr	Label	Example
a.	SK	specific known	<i>Somebody</i> called. Guess who?
b.	SU	specific unknown	I heard <i>something</i> , but I couldn’t tell what it was.
c.	IR	irrealis	You must try <i>somewhere</i> else.
d.	Q	question	Did <i>anybody</i> tell you anything about it?
e.	CA	conditional antecedent	If you see <i>anybody</i> , tell me immediately.
f.	CO	comparative	John is taller than <i>anybody</i> .
g.	IN	indirect negation	I don’t think that <i>anybody</i> knows the answer.
h.	DN	direct negation	John didn’t see <i>anybody</i> .
i.	FC	free choice	You may kiss <i>anybody</i> .

In order for an indefinite to qualify for a function, it must (i) be grammatical in the context the function specifies; and (ii) have the semantics that the function specifies. For example, *any* does not exhibit the specific functions SK/SU because it is ungrammatical in episodic sentences, cf. (9-a); and *some* does not exhibit the comparative function CO because it does not have a universal meaning specified by CO, cf. (9-b).

- (9) a. Somebody /# anybody called.
 b. Berlin is bigger than any /# some Czech city.
 ‘For all Czech cities it holds that Berlin is bigger than they are.’

We have extended Haspelmath’s original map as follows: i) the indirect negation function has been split into an antimorphic (AM) and an anti-additive (AA) function (cf. Zwarts 1998); ii) three new functions have been introduced contiguous to the free choice area, namely the generic function (GEN), the universal free choice (UFC) function, and the indiscriminative function (IND).

(10) *Our extended map*



Motivation for the addition of UFC and GEN came from comparing in more detail the different items cross-linguistically. As reported below, German *irgend*-indefinites do not exhibit GEN or UFC functions, but exhibit a FC function; Italian *uno qualunque* doesn’t exhibit the UFC function, but does exhibit FC and GEN; Dutch *wie*

dan ook doesn't exhibit the GEN function, but exhibits UFC and FC; and finally Italian *qualunque*, Czech *kterýkoli* and Spanish *cualquiera* exhibit all three functions. Motivation for the addition of IND comes from consideration of examples such as the following (from Horn 2005):

(11) I do not want to go to bed with just *anyone* anymore. I have to be attracted to them sexually.

The new functions we added are marked with a \rightarrow in the following illustration:

(12) *Functions on the map*

	Abbr	Label	Example
a.	SK	specific known	<i>Somebody</i> called. Guess who?
b.	SU	specific unknown	I heard <i>something</i> , but I couldn't tell what it was.
c.	IR	irrealis	You must try <i>somewhere</i> else.
d.	Q	question	Did <i>anybody</i> tell you anything about it?
e.	CA	conditional antecedent	If you see <i>anybody</i> , tell me immediately.
f.	CO	comparative	John is taller than <i>anybody</i> .
g.	DN	direct negation	John didn't see <i>anybody</i> .
\rightarrow h.	AM	anti-morphic	I don't think that <i>anybody</i> knows the answer.
\rightarrow i.	AA	anti-additive	The bank avoided taking <i>any</i> decision.
\rightarrow j.	FC	free choice	You may kiss <i>anybody</i> .
\rightarrow k.	UFC	universal free choice	John kissed <i>any</i> woman with red hair.
\rightarrow l.	GEN	generic	<i>Any</i> dog has four legs.
\rightarrow m.	IND	indiscriminative	I don't want to sleep with just <i>anybody</i> anymore.

During annotation we also introduced several off-map functions to label uses which were not strictly indefinite. One example is the *no-matter* function, for which we give here an illustration in Czech:

(13) Ať už jsme v kterékoli zemi, všude nacházíme slušné lidi.
 let already be:1PL in any country everywhere find:1PL polite people
 'No matter in which country you are, you can find polite people everywhere.'

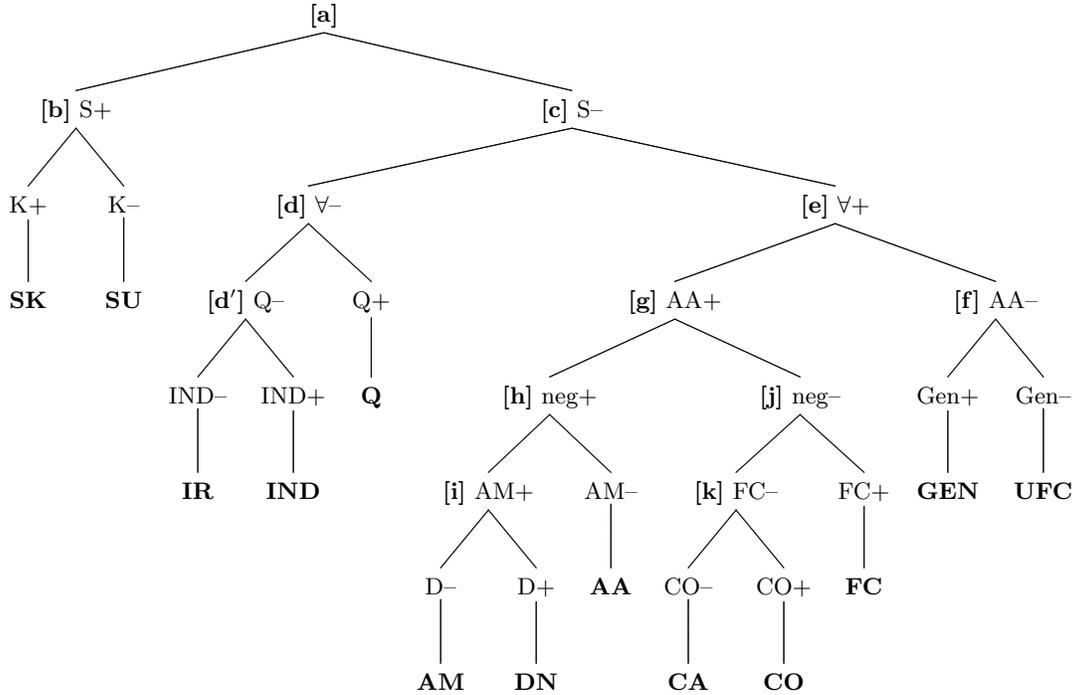
The *no-matter* function played a central role in the diachronic research.

Other functions which were needed in the annotation of individual language data are described in the relevant sections below.

3.2 Methodology for functional annotation

In this section we introduce a set of tests which we used to assign exactly one function to each instance of the examined indefinites. These tests and the order in which they were applied are schematized in the following decision tree. Off-map functions are not represented in this tree.

(14) *Decision tree*



For each node in the decision tree we give now the corresponding test, and, as an illustration, we apply it to the sentences we have used in (12) to exemplify our functional labels.

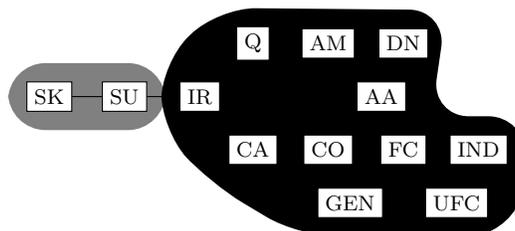
- (a) Test for specificity [S+/-]: Tests whether an indefinite expression refers to a specific individual in the actual world ([S+]) or not ([S-]).

Sentence (S): ...indefinite_i ... **Possible Continuation (PC):** ... pronoun_i ... [S+]

Examples:

- | | | |
|----|--|------|
| a. | <i>Somebody_i</i> called. She _i wanted a new appointment. | [S+] |
| b. | I heard <i>something_i</i> . It _i was very loud. | [S+] |
| c. | You must try <i>somewhere_i</i> else. # It _i is a very nice place. | [S-] |
| d. | Did <i>anybody_i</i> tell you anything about it? # He _i is a real chatterbox. | [S-] |
| e. | If you see <i>anybody_i</i> , tell me immediately. # He _i is a nice guy. | [S-] |
| f. | John is taller than <i>anybody_i</i> . # He _i is short. | [S-] |
| g. | John didn't see <i>anybody_i</i> . # He _i was very tall. | [S-] |
| h. | I don't think that <i>anybody_i</i> knows the answer. # He _i did not even try. | [S-] |
| i. | The bank avoided taking <i>any</i> decision _i . # It _i was difficult. | [S-] |
| j. | You may kiss <i>anybody_i</i> # She _i is beautiful. | [S-] |
| k. | John kissed <i>any</i> woman _i with red hair. # She _i is Italian. | [S-] |
| l. | <i>Any</i> dog _i has four legs. # It _i is very cute. | [S-] |
| m. | I do not want to go to bed with <i>just anyone_i</i> anymore. # He _i is very handsome. [S-] | [S-] |

The application of test (a) splits our map into a specific area (in grey) and a non-specific area (in black).



Within the specific area we apply test (b) to distinguish the specific known from the specific unknown function.

- (b) Test for known [K+/-]: Tests whether the speaker knows who the referent of an indefinite is ([K+]) or not ([K-]).

S:...indefinite ... **PC**: Guess who/what?

[K+]

Examples:

a. *Somebody* called. Guess who?

[K+] \mapsto [SK]

b. I heard *something*, but I couldn't tell what it was. # Guess what?

[K-] \mapsto [SU]

Within the non-specific area we apply test (c) to distinguish between wide-scope universal meaning and genuinely existential meaning:

- (c) Test for universal meaning [V+/-]:

... **Op** (...indefinite ...) ... \Rightarrow ... $\forall x$ (**Op**... x ...) ...

Examples:

a. You must try *somewhere* else \nRightarrow for all places x : you must try x

[V-]

b. Did *anybody* tell you anything about it? \nRightarrow for every x : did x tell you about it?

[V-]

c. If you see *anybody*, tell me immediately \Rightarrow for every x : if you see x , tell me immed.

[V+]

d. John is taller than *anybody* \Rightarrow for every x : John is taller than x

[V+]

e. I didn't see *anybody* \Rightarrow for every x : I didn't see x

[V+]

f. I don't think that *anybody* knows the answer \Rightarrow for every x : I don't think x knows the answer

[V+]

g. The bank avoided taking *any* decision \Rightarrow for every decision x : the bank avoided taking x

[V+]

h. You may kiss *anybody* \Rightarrow for every x : you may kiss x

[V+]

i. John kissed *any* woman with red hair. \Rightarrow for every woman x with red hair: John kissed x

[V+]

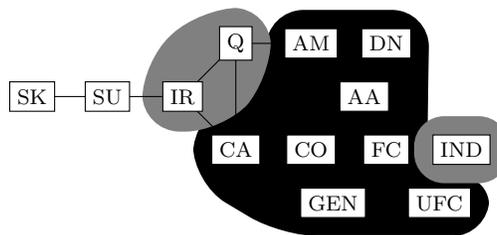
j. *Any* dog has four legs \Rightarrow for every dog x (with exceptions?): x has four legs

[V+]

k. I do not want to go to bed with *just anyone* _{i} anymore. \nRightarrow for all x : I do not want to go to bed with just x anymore

[V-]

The application of test (c) splits the non-specific area into an existential area (in grey) and a wide-scope universal area (in black).



Within the existential area we distinguish polar questions from irrealis non-specific and indiscriminative uses via step (d).

- (d) Test for polar question [Q+/-]: ...indefinite _{i} ... **PC**: ... or not?

[Q+]

Examples:

a. You must try *somewhere* else (# or not?)

[Q-]

b. I do not want to go to bed with *just anyone* _{i} anymore (# or not?)

[Q-]

c. Did you see *anybody* (or not)?

[Q+] \mapsto [Q]

Also within the existential area, we apply test (d') to distinguish irrealis non-specific from indiscriminative uses.

- (d') Test for indiscriminacy [IND+/-]: Tests for free choice (or randomness) flavor via compatibility with *just*.

Examples:

a. You must try (#just) *somewhere* else.

[IND-] \mapsto [IR]

b. I do not want to go to bed with *just anyone* _{i} anymore

[IND+] \mapsto [IND]

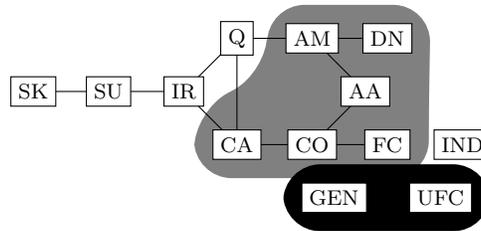
Within the wide-scope universal area we apply test (e) to distinguish anti-additive contexts from non anti-additive ones.

(e) Test for anti-additivity [AA+/-]: $\mathbf{Op}(a \vee b) \Rightarrow \mathbf{Op}(a) \wedge \mathbf{Op}(b)$ [AA+]

Examples:

- a. If you see *anybody*, you should tell me immediately. [If you see John or Mary, you should tell me immediately \Rightarrow If you see John, you should tell me immediately and if you see Maria, you should tell me immediately] [AA+]
- b. John is taller than *anybody*. [John is taller than Lee or Mary \Rightarrow John is taller than Lee and John is taller than Mary] [AA+]
- c. John didn't see *anybody*. [John didn't see Lee or Mary \Rightarrow John didn't see Lee and John didn't see Mary] [AA+]
- d. I don't think that *anybody* knows the answer. [I don't think that Mary or Lee know the answer \Rightarrow I don't think that Mary knows the answer and I don't think that Lee knows the answer] [AA+]
- e. The bank avoided taking *any* decision. [The bank avoided taking decision A or decision B \Rightarrow The bank avoided taking decision A and the bank avoided taking decision B] [AA+]
- f. You may kiss *anybody*. [You may kiss John or Mary \Rightarrow you may kiss John and you may kiss Mary] [AA+]
- g. John kissed *any* woman with red hair. [John kissed Lee or Bea \nRightarrow John kissed Lee and John kissed Bea] [AA-]
- h. *Any* dog has four legs. [Fido or Bobby has four legs \nRightarrow Fido has four legs and Bobby has four legs] [AA-]

The application of test (e) splits the universal area into an anti-additive area (in grey) and a non anti-additive area (in black).



Within the non anti-additive area we apply test (f) to distinguish generic from universal free choice readings.

(f) Test for genericity [Gen+/-]: ...indefinite ... \equiv ...plain generic indef. ... [Gen+]

Examples:

- a. John kissed *any* woman with red hair \nRightarrow John kissed a woman with red hair [Gen-] \mapsto [UFC]
- b. *Any* dog has four legs \equiv A dog has four legs [Gen+] \mapsto [GEN]

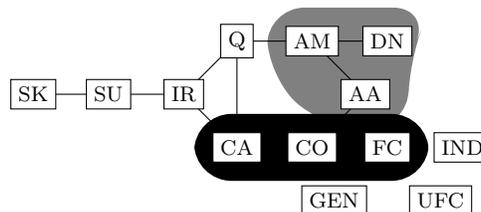
Within the anti-additive area we apply test (g) to distinguish negative contexts from non negative ones.

(g) Test for negative meaning [Neg+/-]: $\mathbf{Op}(a \vee \neg a)$ is inconsistent [Neg+]

Examples:

- a. John didn't see *anybody*. [John didn't stay or go \mapsto inconsistent] [Neg+]
- b. I don't think that *anybody* knows the answer. [I don't think that the door is open or closed \mapsto inconsistent] [Neg+]
- c. The bank avoided taking *any* decision. [The bank avoided being open or closed] \mapsto inconsistent [Neg+]
- d. You may kiss *anybody*. [You may stay or go \mapsto not inconsistent] [Neg-]
- e. If you see *anybody*, you should tell me. [If you stay or go, you should tell me \mapsto not inconsistent] [Neg-]
- f. John is taller than *anybody*. [John is taller than somebody or nobody \mapsto not inconsistent] [Neg-]

The application of test (g) splits the anti-additive area into a negative area (in grey) and a non-negative area (in black).



Within the negative area we apply test (h) to distinguish anti-multiplicative contexts from plain negative ones.

(h) Test for anti-multiplicativity: $\mathbf{Op}(a) \vee \mathbf{Op}(b) \equiv \mathbf{Op}(a \wedge b)$ [AM+]

Examples:

- a. John didn't see *anybody*. [John didn't see Mary or John didn't see Sue \equiv John didn't see (Mary and Sue)] [AM+] [AM+]
- b. I don't think that *anybody* knows the answer. [I don't think that Lee knows the answer or I don't think that Mary knows the answer \equiv I don't think that (Lee and Mary) know the answer] [AM+] [AM+]
- c. The bank avoided taking *any* decision. [The bank avoided taking decision A or the bank avoided taking decision B \neq The bank avoided taking (decision A and decision B)] [AM-] \mapsto [AA]

Within the anti-multiplicative area we check if the relevant operator is clausal negation.

- (i) Test for direct negation [D+/-]: **Op** is clausal negation [D+]

Examples:

- a. John didn't see *anybody*. [D+] \mapsto [DN]
- b. I don't think that *anybody* knows the answer. [D-] \mapsto [AM]

Within the anti-additive non negative area we apply test (j) to distinguish free choice contexts.

- (j) Test for free choice [FC+/-]: **Op**($a \vee \neg a$) is informative [FC+]

Examples:

- a. If you see *anybody*, you should tell me. [If you stay or go, you should tell me \mapsto antecedent is not informative] [FC-]
- b. John is taller than *anybody* [John is taller than somebody or nobody \mapsto not informative] [FC-]
- c. You may kiss *anybody* [You may stay or go \mapsto informative] [FC+] \mapsto [FC]

Within the non free choice contexts we distinguish the comparative constructions from the others.

- (k) Test for comparative construction [CO+/-]: Indefinite occurs in comparative construction [CO+]

Examples:

- a. If you see *anybody*, tell me immediately. [CO-] \mapsto [CA]
- b. John is taller than *anybody*. [CO+] \mapsto [CO]

When tests (a)-(k) failed to yield a conclusive result, the example was classified as "Unclear".

Further applications of these tests, including their application to ambiguous examples, are discussed in detail in Aguilar-Guevara et al. (2011).

4 The database

The Corpus of Indefinite Uses is available via an online web interface at the following url: indefdb.logophile.org. To gain access to the online database, contact Maria Aloni (M.D.Aloni@uva.nl).

After logging on to the database, you will be directed to the Main Menu, where you will have the choice of several predefined searches by language, or of conducting a custom search. Clicking on the tab "Custom Search" on the main screen will lead to the Search Screen (Fig. 1). Here, the database can be searched by field and by free text.

Examples in the database have been annotated with the following tags, all of which are searchable.

- Item: Overall term for the indefinite item or series (examples: German *irgend**; Italian *qualunque*)
- Form: Subclassification of item, used when an indefinite occurs in two or more distinct forms (examples: German *irgendein* and *irgend jemand*; Italian plain *qualunque*, *un N qualunque*, *un qualunque N*)
- Source: Source of original corpus example (when available)
- Language
- Year: Year of original corpus example (when available)
- Function: Functional semantic classification on the expanded Haspelmath map, as described in Section 3 above
- Semantic tag: Optional tag providing information about the semantics of the indefinite form

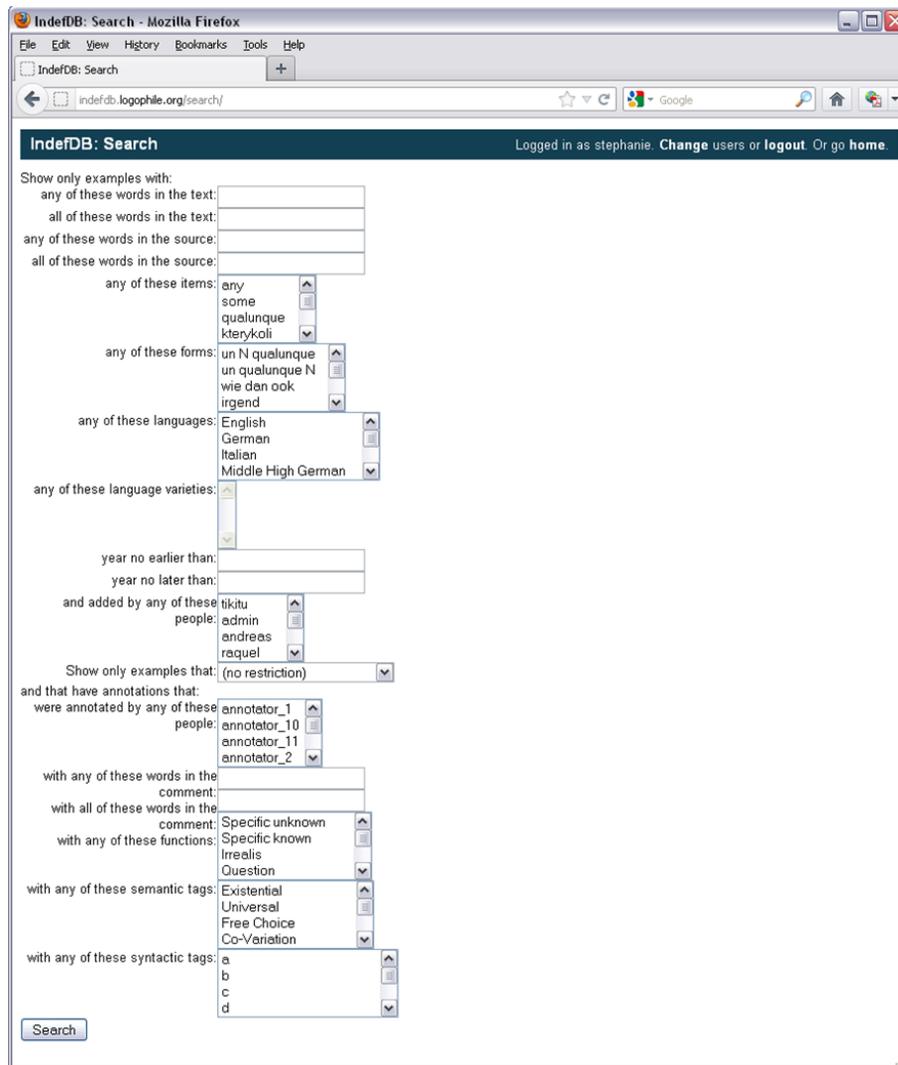


Figure 1: Search screen

- Syntactic tag: Optional tag providing information about the syntactic status of the indefinite form and/or the syntactic environment in which it occurs
 - NB: To search by function, semantic tag or syntactic tag, first go to the field ‘Show only items that’ and highlight the option ‘have been annotated’.

From the results screen, there is also the option to export search results to a text file, which can be imported into another package (e.g. Excel).

As the nature of the annotation varied by language, further information on the tags used is given in the individual language sections below.

5 Individual language data and results: synchronic

The searchable database includes corpus examples of indefinite forms from six languages, which formed the basis of a cross-linguistic corpus study: German *irgend* series, Czech *kterýkoli*, Spanish *cualquiera*, Italian (*uno*) *qualunque*, Dutch *wie dan ook* and English *some* and *any*. In this section we describe the synchronic data available for each language and how it has been annotated, and present the most significant findings of the corpus study. The diachronic data is described in the following section.

5.1 German

Data

- Indefinite forms included: *irgend* series, including *irgendein* [*irgend* + *ein* ‘a’] and *irgend jemand* - [*irgend* + *jemand* ‘someone’]
- Corpus: DWDS (Berlin-Brandenburgische Akademie der Wissenschaften; 100 million tokens, written, various registers including fiction, newspapers, science and non-fiction)
- Queries: *irgendein** [six possible forms: *irgendein*, *irgendeine*, *irgendeiner*, *irgendeines*, *irgendeinen*, *irgendeinem*]; *Irgendein**; *irgendjemand**; *irgend jemand**; *Irgendjemand**; *Irgend jemand**
- Date of search: June 2008
- Number of occurrences: *irgendein* - 5975, out of which 5045 available (due to copyright); *irgend jemand* - 823, out of which 610 available
- Labeled: 300 random occurrences of each *irgendein* and *irgend jemand*

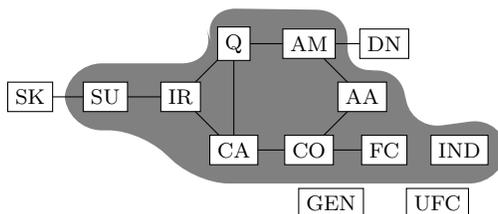
Annotation

- Item: *irgend**
- Forms: *irgendein*; *irgend jemand*
- Functions: Those described in Section 3 only
- Syntactic tag:
 - In combination with Free Choice function, to indicate syntactic context in which indefinite form occurs: modal possibility, MP[modal possibility]-deontic, MN[modal necessity]-deontic, MP-epistemic, M[modal]-teleological, M-ability, M-bulethic
 - In combination with Indiscriminative function, to indicate syntactic context: anti-morphic
 - In combination with Irrealis function, to indicate syntactic context: MP-epistemic, MN-epistemic
 - In combination with Conditional Antecedent function, to mark additional contexts classified as CA: Restrictor, Superlative
- Semantic tag:
 - Used when example is potentially ambiguous: free choice, indifference, ignorance, non-specific, co-variation
- Source: provided
- Year: provided

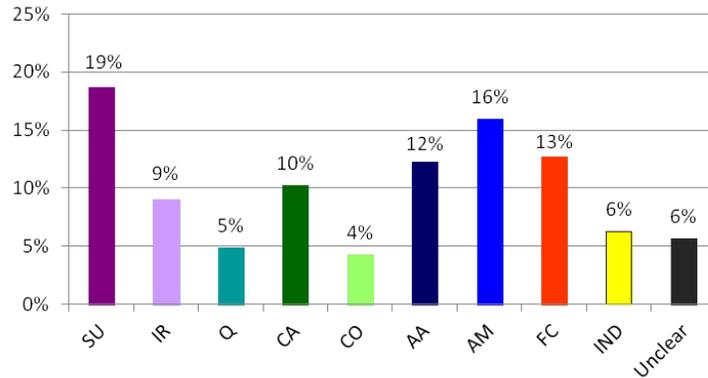
Results

Irgend indefinites occur in the functions depicted below:

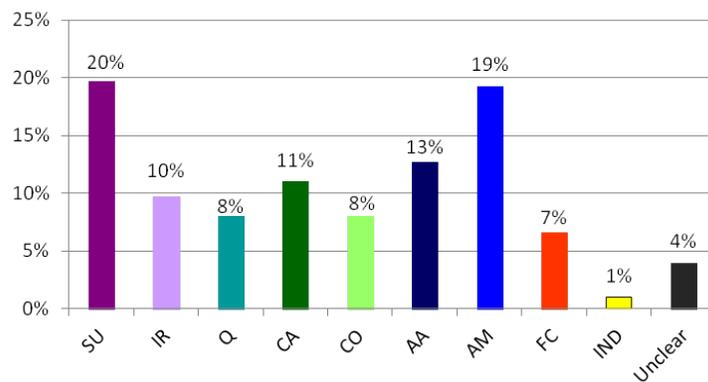
(15) *Distribution*



Irgendein:



Irgend jemand



Typical examples from the corpus of each of the attested functions are given below:

- (16) Irgend jemand versetzte dem Gefangenen mit einem Stock einen so schweren Schlag über den Kopf, daß er eine stark blutende Wunde davontrug. [SU]
 IRGEND-SOMEBODY gave DEF-DAT prisoner with INDEF-DAT stick INDEF-ACC so hard blow over DEF-ACC head, that he INDEF-ACC intense bleeding wound carried-away.
 ‘Somebody gave the prisoner such a heavy blow on the head with a stick, that he came away with a heavily bleeding wound.’
- (17) Lisa vermutete, daß Herbert irgendeinen amtlichen Brief geschrieben haben wollte. [IR]
 Lisa suspected that Herbert IRGEND-ONE official letter written have wanted
 ‘Lisa suspected that Herbert wanted to have some official letter written.’
- (18) Hast Du irgendeine Tageszeitung? [Q]
 Have you IRGEND-ONE newspaper?
 ‘Do you have any newspaper?’
- (19) ... wenn Sie irgendeinen Wunsch haben, Kaffee und so weiter, rufen Sie einfach unten an. [CA]
 ... if you IRGEND-ONE wish have, coffee and so on, call you just downstairs PRT.
 ‘...if you want anything, coffee and so on, just call downstairs’
- (20) Er besaß zehn Milliarden Dollar, mehr als irgendein Mensch auf der Welt. [CO]
 He owned ten billion dollars, more than IRGEND-ONE person on the world.
 ‘He had ten billion dollars, more than any person in the world’
- (21) Ich glaube nicht, daß sie irgendetwas bewirken werden oder daß irgend jemand sie glaubwürdig findet. [AM]
 I believe not, that they anything effect will or that IRGEND-SOMEBODY they credible find.
 ‘I don’t think that they will achieve anything or that anybody will find them credible’

- (22) Es ist nie irgend jemand zu euch gekommen mit einer solchen Botschaft wie wir: Jesus Christus ist auferstanden! [AA]
 It is never IRGEND-SOMEBODY to you come with a such message like us: Jesus Christ is risen!
 ‘Never has anybody come to you with such a message like us: Jesus Christ is risen!’
- (23) [...] du begleitest sie natürlich, um dort bei deiner Mutter das wahre Verlobungsfest zu feiern und [...] you escort her of course, to there at your mother that true engagement to celebrate and schreibst nach einiger Zeit einen kühlen Absagebrief, du kannst ja auch irgendein Geschenk hinzufügen, write after some time a cool rejection letter, you can prt also IRGEND-ONE present add, was bei solchen Mädchen immer Eindruck macht. [FC]
 which for such girls always impression makes
 ‘...of course you must escort her to celebrate the real engagement at your mother’s house and after some time you write a cool rejection letter, you can also add a present, (anything really). That will always impress such girls.’
- (24) Das ist wenn man Lust auf einen Mann hat, Madame! Nich auf irgendeinen. Auf den! [IND]
 That is when one desire for INDEF-ACC man has, madam! Not for IRGEND-ONE. For DEF-ACC!
 ‘That’s when you desire a man, Madam, not just any man, but HIM !’

5.2 Czech

Data

- Indefinite forms included: *kterýkoli* [*který* ‘which’ + *koli*; *li* is a particle that can mark (in the form of a verbal affix) embedded yes/no questions and conditional clauses, though only in a formal register]
- Corpus: Český národní korpus (ČNK) (Czech national corpus); Subcorpus SYN; URL <http://korpus.cz/corpora/>
- Query: Kterýkoli in all its possible forms: 6 grammatical cases (nom, acc, gen, dat, loc, instr) + 6 paradigms (sg.masc.anim, sg.masc.inanim, sg.fem, sg.neut, pl.masc.anim, pl.rest); that gives 11 forms (due to syncretisms) times 2 (initial letter small/capital); *kterýkoli*, *kteréhokoli*, *kterémukoli*, *kterémkoli*, *kterýmkoli*, *kterékoli*, *kterákoli*, *kteroukoli*, *kteríkoli*, *kterýchkoli*, *kterýmikoli*, *Kterýkoli*, *Kteréhokoli*, *Kterémukoli*, *Kterémkoli*, *Kterýmkoli*, *Kterékoli*, *Kterákoli*, *Kteroukoli*, *Kteríkoli*, *Kterýchkoli*, *Kterýmikoli*
- Date of search: 29-10-2009
- Number of occurrences: 7843
- Labeled: 300 random occurrences

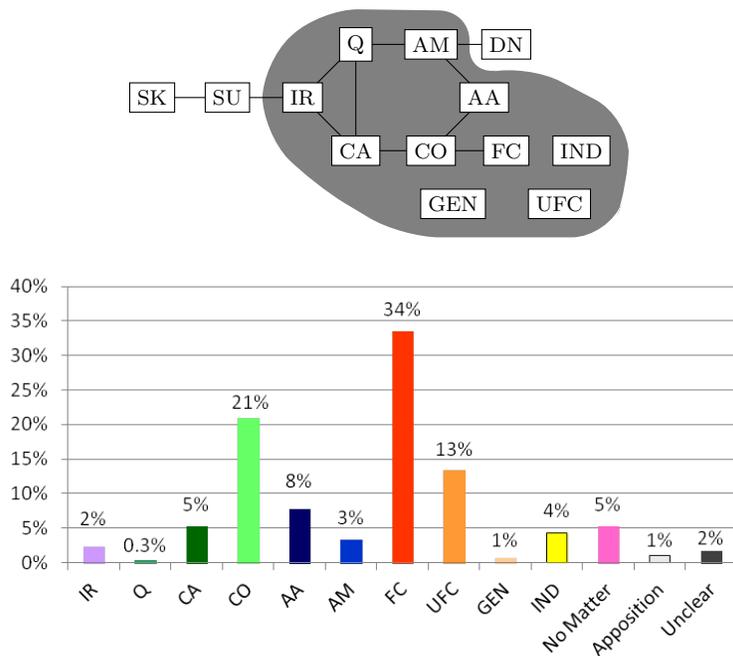
Annotation

- Item: *kterýkoli* [appears without diacritics]
- Forms: none
- Functions: Those described in Section 3 (including No Matter) + Apposition
- Syntactic tag:
 - As subcategory of Irrealis function: Modal, MN-deontic, sufficient-condition, M-teleological
 - As subcategory of Conditional Antecedent function: Antecedent-conditional
 - As subcategory of Anti-additive function: Post-nominal modification, restrictor
 - As subcategory of Free Choice function: Modal, sufficient-condition, imperative, MP-deontic, MP-epistemic, M-ability
 - As subcategory of Universal Free Choice function: Modal, MN-deontic, veridical-habitual, postnominal modification
 - As subcategory of Generic function: Veridical-habitual, Veridical-generic
 - As subcategory of Indiscriminative function: Modal, Veridical-episodic

- Semantic tag:
 - To clarify the interpretation of the indefinite form: existential, universal, free choice, indifference
- Source: not available
- Year: not available

Results

Czech *kterýkoli* has the distribution illustrated below.



Typical examples from the corpus of each of the attested functions are given below:

- (25) Také jsem mu poručil, aby odvedl Mehunefer na kteroukoli do Vesetu mířící loď se vší nezbytnou zdvořilostí – a bude-li nutno, i násilím. [IR]
 also AUX.PAST.1SG him ordered that take Mehunefer on any to Veset heading ship with all necessary politeness and will.be-COND necessary even violence.INSTR
 ‘I also ordered him to take Mehunefer to any ship heading towards Veset with all the politeness needed – and if necessary, even using violence.’
- (26) Otázka je, zdali je kterýkoli z nich natolik významný, abychom kvůli němu dělali zásadní teoretické rozlišení mezi tím, co... [Q]
 question is whether is ANY of them so important that.1PL because.of him do crucial theoretical distinction between that what
 ‘The question is whether any of them is so important so that because of that we make a crucial theoretical distinction between what...’
- (27) Je-li podle nich vůz uznán v kterékoli členské zemi, je uznán všude... [CA]
 is-COND according.to them car acknowledged in any member state is acknowledged everywhere
 ‘If a car is acknowledged in any member state, it is acknowledged everywhere...’
- (28) A tak se v šesti letech stal z Nata pokrytec. Byl o pět centimetrů vyšší než kterýkoli z jeho trapičů. [CO]
 and so REFL in six years became from Nat hypocrite was of five centimeters taller than any of his teasers
 ‘And so Nat became a hypocrite when he was six. He was five centimeters taller than any one of his teasers.’

- (29) Není moudré, když si kterákoli ze stran klade podmínky, aby se obnovilo jednání, ... [AM]
 NEG.IS wise when REFL any from parties puts conditions that REFL renew negotiation
 ‘It’s not wise for any of the parties to dictate conditions for a renewal of the negotiation, ...’
- (30) Z 11 jich prý ODS požaduje sedm a se třemi počítá pro ODA. Odmítli jsme absolutní
 from 11 them allegedly ODS asks seven and with three counts for ODA refuse AUX.PAST.1PL absolute
 většinu kterékoli ze stran, uvedl Malíř. [AA]
 majority any from parties said Malíř
 ‘Allegedly, ODS asks for seven out of eleven and they count with three for ODA. We refused the absolute majority of any of the parties, said Malíř.’
- (31) Musíte říkat gej,” opravil Bendu jeden z přítomných a vysvětlil: “Vysloveno gáj, to je anglicky
 must.2 say [gay] corrected Benda.ACC one of present and explained pronounced [guy] that is English
kterýkoli chlapec. Tedy i vy.” [IND]
 any boy so even you
 ‘You have to say gay,” one of the people present corrected Benda and explained: “If you say guy, that means any boy in English. That is, even you.’
- (32) Z hlediska leteckých a raketových úderů je zaměřitelný kterýkoli objekt a kterýkoli objekt je také
 from viewpoint air and rocket attacks is targetable any object and any object is also
 teoreticky zničitelný, ... [FC]
 theoretically destroyable
 ‘Seen from the viewpoint of air and rocket attacks any object can be targeted and theoretically, any object can also be destroyed, ...’
- (33) ... to, že při návratu z nákupu najde své auto na místě, nese provozovatel parkoviště.
 that.DEM that.COMPL at return from shopping finds his.REFL car in place bears operator parking.lot
 Tu ostatně nese kterýkoli provozovatel, v kterémkoli městě. [GEN]
 that by.the.way bears any operator in any town
 ‘It is the operator of the parking lot who bears the responsibility that [somebody] finds his car in place at the return from shopping. By the way, any operator bears that responsibility, in any town.’
- (34) Já v tom ale žádný podraz nevidím: ta možnost je přece zakotvena v ústavě a je svaté
 I in it but any.NCI four.trick NEG.SEE.1SG the possibility is particle anchored in constitution and is holy
 právo kteréhokoli kandidáta na ni upozornit. [UFC]
 right any candidate.GEN to it draw.attention
 ‘But I don’t see any foul trick in it: one should keep in mind that the possibility is anchored in the constitution and it is the holy right of any candidate to draw attention to it.’
- (35) Takovýhle návrh od Gorbovského rozhodně nečekal. Bylo to divoké, ať se to vzalo z
 such proposal from Gorbovsky definitely NEG.EXPECTED was it wild PARTICLE REFL it take from
kterékoli strany. [No matter]
 any perspective
 ‘He definitely didn’t expect such a proposal from Gorbovsky. It was wild, no matter from which perspective one looked at it.’
- (36) Myslím, že to tak není. Když k nám přijede soupeř, a to kterýkoli, s podobnou taktikou jako
 think that it so NEG.IS when to us comes opponent and that any with similar tactics like
 slávisté v neděli, tak se nám hraje pochopitelně velmi špatně. [Apposition]
 Slavia on Sunday then REFL us play of.course very badly
 ‘I don’t think it’s that. When an opponent, any [opponent], is coming with a similar tactics like Slavia had on Sunday, then it’s clear we have a bad time playing.’

5.3 Spanish

Data

- Indefinite form included: *cualquier(a)* [*cual-* ‘which’ + *quiera* ‘want:PRES.SUBJ.3’]
- Corpus: CORPUS DEL ESPAÑOL (by Mark Davies; 100 million words; spoken, fiction, newspaper and academic texts; 13th-20th centuries)
- Query: *ualq* [all possible forms of *cualquier(a)* + 10 instances of completely unrelated words, which were

excluded]; restricted to 20th century, spoken

- Date of search: 2009
- Number of occurrences: 7744
- Labeled: 200 random occurrences. NB: The CORPUS DEL ESPAÑOL makes available for downloading only a brief (ca. 15-25 word) context for each token. Expanded contexts are available through the corpus' online interface, and were consulted during the annotation process.

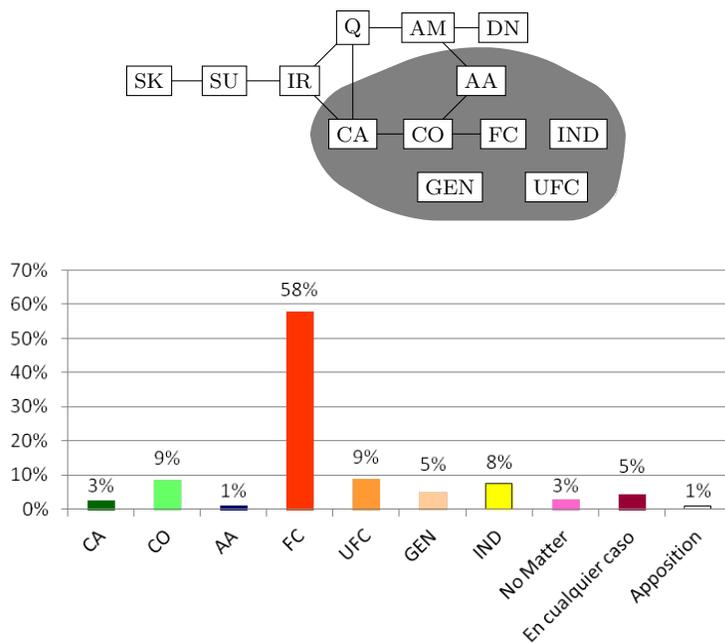
Annotation

- Item: cualquiera
- Forms: none
- Functions: Those described in Section 3, including No matter (Spanish *cualquiera que sea*) plus En cualquier caso ('in any case') and Apposition
- Syntactic tag: None used
- Semantic tag: None used
- Source: Provided
- Date: Only century available; synchronic data annotated as '20' (for 20th century)

Results

In modern spoken Spanish, *cualquier(a)* has the distribution illustrated below. As seen here, the FC use dominates, while contiguous functions on the map are also represented. Three off-the-map functions are also found in the corpus: No matter, en cualquier caso (in any case) and apposition.

(37) *Distribution*



For each of the attested functions, a typical corpus example is given below:

- (38) Rigoletto es un camarero que podría encontrarse en cualquier local de Madrid. [FC]
 Rigoletto be:3SG a waiter who could find:INF.REFL in CUALQUIER establishment of Madrid
 ‘Rigoletto is a waiter who could be found in any establishment in Madrid.’
- (39) Tuvieron que aceptar, para subsistir, cualquier trabajo que se les ofrecía. [UFC]
 have:PST.3PL to accept:INF, to subsist:INF, CUALQUIER work that REFL them offer:PST.3SG
 ‘They had to accept, to subsist, any work that was offered to them.’
- (40) Cualquier acontecimiento artístico necesita del público. [GEN]
 CUALQUIER event artistic need:3SG of.the audience
 ‘Any artistic event is in need of an audience.’
- (41) cualquier película no me convence, sino tiene que ser algún tipo especial de películas. [IND]
 CUALQUIER film not me convince:3SG, but have:3SG to be:INF some type special of films
 ‘Not just any film convinces me - it has to be a special type of film.’
- (42) Un cantante está expuesto a todo eso mucho más que cualquier otro músico. [CO]
 A singer be:3SG expose:PP to all that much more than CUALQUIER other musician
 ‘A singer is exposed to all that much more than any other musician.’
- (43) Si estás haciendo cualquier tipo de trabajo de investigación ... [CA]
 If be:2SG do:PROG CUALQUIER type of work of investigation ...
 ‘If you are doing any type of investigative work ...’
- (44) Está prohibido que cualquier persona o joven o niño emplee caucheras para tumbar a los pajaritos. [AA]
 Be:3SG prohibit:PP that CUALQUIER person or youth or child employ:3SG.SUBJ slingshots to knock.down to the birdies
 It is prohibited for any person or youth or child to use slingshots to knock down the little birds.
- (45) Pero algunos personajes permanecían a flote, estaban siempre en el cargo, cualquiera que fuera el Gobierno. [No matter]
 be:PST.SUBJ.3PL the government
 ‘But some important people remained afloat, were always in office, no matter who the government was’
- (46) En cualquier caso es, en mi opinión, injusto. [En cualquier caso]
 In cualquier case be:3SG, in my opinion, unjust
 ‘In any case it is, in my opinion, unjust’

5.4 Italian

Data

- Indefinite form included: *qualunque* [from Latin *qualiscúmque* composed from *qualis* ‘what’ + *cúmque* ‘ever’]
- Corpus: CORIS (100 million words; various registers; a synchronic corpus of written language, whose component texts belong, roughly speaking, to the 1980s and 1990s, with a somewhat wider temporal collocation as far as narrative is concerned)
- Date of search: 13-8-2008
- Number of occurrences: 7591
- Labeled: 300 random occurrences

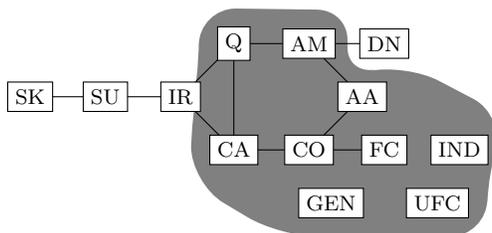
Annotation

- Item: qualunque
- Forms: plain qualunque, un qualunque N, un N qualunque (where ‘un’ stands for any indefinite determiner), Det N qualunque (where Det is not an indefinite determiner), other

- Functions: Those described in Section 3 only
- Syntactic tag:
 - In combination with Indiscriminative function, to indicate syntactic context of indefinite form: Direct-negation, veridical-episodic, veridical-generic, veridical-habitual, veridical-copula, comparative, conditional, modal
 - In combination with Free Choice function, to indicate syntactic context: Imperative, modal, sufficient condition
 - In combination with Question function, to mark syntactic context: modal
 - In combination with Anti-Additive, to mark context classified as AA: before
 - In combination with Conditional Antecedent, to mark context classified as CA: restrictor
- Semantic tag:
 - average, random, low level (subcategories of Indiscriminative function)
 - Domain-widening (when strong widening effect is clear)
- Source: Provided
- Date: Not available

Results

Overall, Italian *qualunque* has the distribution shown below:



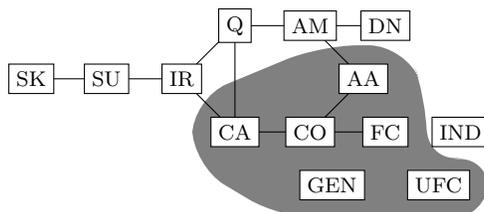
- *Qualunque* can occur in these forms:

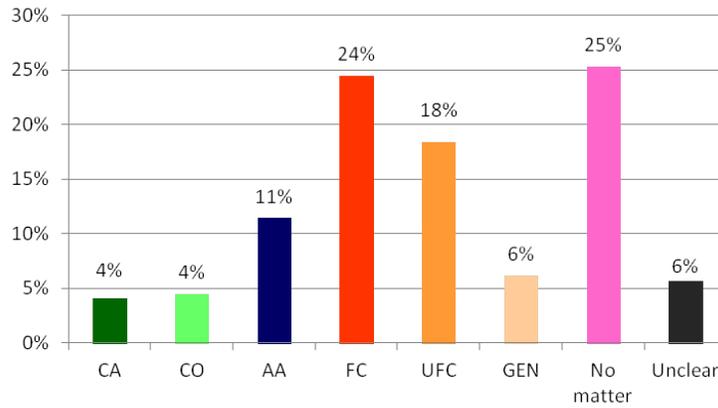
(47)	a.	Plain <i>qualunque</i>	245 (81,7%)
	b.	Un <i>qualunque</i> N	22 (7,3%)
	c.	un N <i>qualunque</i> ('un' stands for any indefinite determiner)	25 (8,3%)
	d.	Det N <i>qualunque</i> (Det not indefinite determiner)	7 (2,3%)
	e.	Other	1 (0,3%)

The distribution of each (as found in the corpus), as well as corresponding examples, are provided below:

Plain *qualunque*

(48) *Distribution*

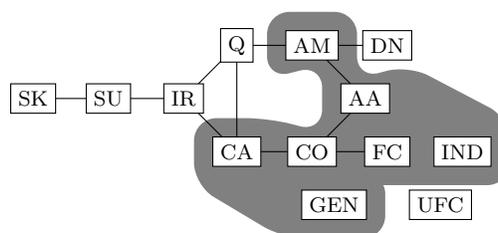


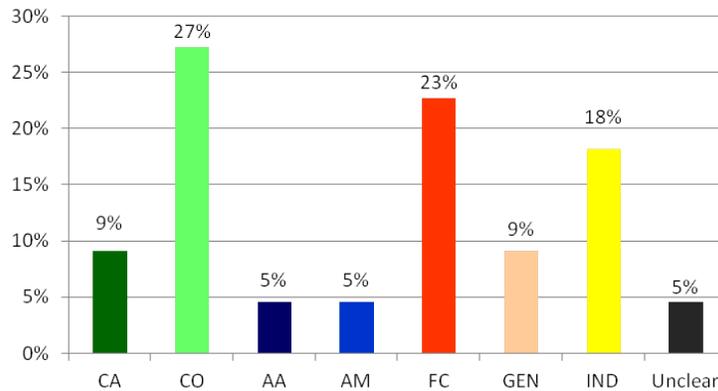


- (49) Irraggiungibile da qualunque tenerezza. Troppo crudele per essere amato. [AA]
 Unreachable by QUALUNQUE tenderness. Too cruel to be loved
 ‘Unreachable by any tenderness. Too cruel to be loved.’
- (50) Guarda il mondo: è più fantastico di qualunque sogno studiato e prodotto dalle più grandi fabbriche [CO]
 Look the world: is more fantastic of QUALUNQUE dream studied and produced from-the most big factories
 ‘Look at the world: it is more fantastic than any dream studied and produced by the biggest factories’
- (51) ogni giocata elementare di qualunque metodo fa storia a sè, indipendentemente da quanto è successo prima [CA - restriction of universal]
 every game elementary of QUALUNQUE methods does history to itself, independently from how-much is happened before
 ‘every basic game of any method has its own history, independently of what happened before’
- (52) Non potevo più sottovalutarli. Erano in grado di ottenere qualunque risultato. [FC]
 Not I-could more underestimate-them. They-were in able to obtain qualunque result
 ‘I could no longer underestimate them. They were able to obtain any result.’
- (53) Lei ha ragione: qualunque dibattito, oggi, si riduce a uno ‘scontro Coppi-Bartali’ [GEN]
 You have reason: QUALUNQUE debate, nowadays, itself reduces to a ‘contest Coppi-Bartali’
 ‘You are right: any debate, nowadays, comes down to a ‘Coppi-Bartali contest’
- (54) Si distruggono tutte le tracce dell’ esistenza del defunto, i suoi arnesi, la sua capanna, qualunque cosa gli sia appartenuta.[UFC]
 One destroys all the traces of-the existence of-the deceased, the his tools, the his hut, QUALUNQUE thing to-him is_{sub} belonged
 ‘One destroys all traces of the existence of the deceased, his tools, his hut, any thing that has belonged to him.’
- (55) Qualunque sia stato il movente, il furto ha comunque fruttato un ricco bottino. [No matter]
QUALUNQUE is_{sub} been the motive, the robbery has anyway produced a rich booty
 ‘Whatever was the motive, the robbery produced a rich booty.’

Un + qualunque + N

- (56) *Distribution*

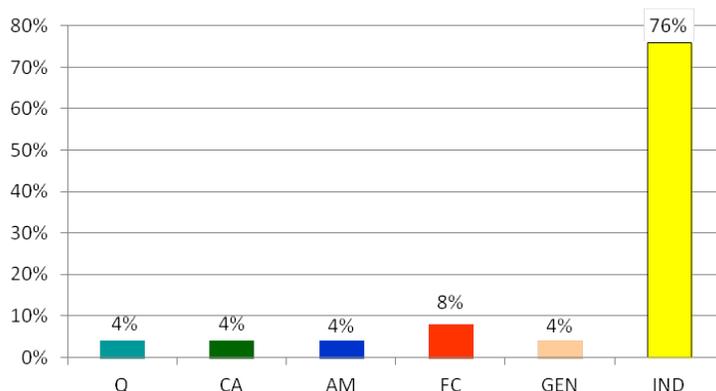
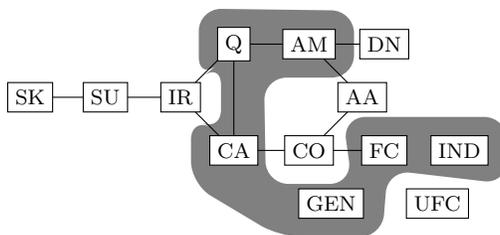




- **Note:** According to native speaker judgments, *un qualunque N* is also possible in the Q function. While no examples of this were identified in the corpus, we attribute this to the small sample size for this form (22 tokens).

- (57) Non ho mai prestato servizio militare nè sono mai stato assimilato a un qualunque servizio militare. [AM]
 Not I-have never done service military nor I-am never been assimilated to a QUALUNQUE service military
 'I never did military service, nor have I ever been assimilated to any military service.'
- (58) Prima che una qualunque delle vicende giudiziarie che hanno coinvolto il fondatore della Fininvest [AA]...
 Before that a QUALUNQUE of-the incidents legal that have embroiled the founder of-the Fininvest
 ...
 'Before any of the legal incidents that have embroiled the founder of the Fininvest ...'
- (59) Se trovate una qualunque incrinatura, non importa se piccola o grande, dovrete effettuare una ... [CA]
 If you-find a QUALUNQUE crack, not matter if small or big, you-will-have-to make a ...
 'If you find any crack, no matter whether small or big, you will have to make a ...'
- (60) un bambino che ha già sofferto non ha bisogno di una qualunque altra famiglia, ma di genitori in grado di [IND]
 a child who has already suffered not has need of a QUALUNQUE other family, but of parents in able to
 rispondere alle sue peculiarità
 answer to-the his ideosyncracies
 'a child who has already suffered doesn't need just any other family, but parents who are able to respond to his ideosyncracies'
- (61) Ad esempio, quando qualcuno è assetato, può cercare una qualunque bevanda che lo disseti. [FC]
 To example, when someone is thirsty, can search-for a QUALUNQUE beverage that him refreshes_{sub}
 'For example, when someone is thirsty, he can search for any beverage which will refresh him'
- (62) e non sono più dannosi di un qualunque impianto industriale [CO]
 and not they-are more harmful of a QUALUNQUE establishment industrial
 'and they are not more harmful than any industrial establishment'
- (63) Curva. L' insieme dei punti costituenti una qualunque linea non retta. [GEN]
 Curve. The set of-the points constituting a QUALUNQUE line not straight.
 'Curve. The set of points constituting any non straight line'

Un + N + qualunque



- Note: According to native speaker judgments, *un N qualunque* is also possible in the CO and AA functions. While no examples of these were identified in the corpus, we attribute this to the small sample size for this form (25 tokens).

- (64) Ci può suggerire una spiegazione qualunque? [Q]
 to-us you-can suggest a explanation QUALUNQUE?
 ‘Can you suggest to us any explanation?’
- (65) Per lo psicoanalista ciò significa che un cambiamento in una parte qualunque della vita psichica sconvolge
 For the psychoanalyst this means that a change in a part QUALUNQUE of-the life psychic devastates
 l’ equilibrio raggiunto in precedenza [CA-restriction]
 the balance reached previously
 ‘For the psychoanalyst this means that a change in any part of the psychic life devastates the previously reached balance.’
- (66) Il gruppo può essere composto da un numero qualunque di documenti, ma i documenti devono trovarsi
 The group can be composed by a number QUALUNQUE of files, but the files must be
 tutti nella stessa cartella. [FC]
 all in-the same folder.
 ‘The group can be composed by any number of files, but the files must all be in the same folder.’
- (67) Hanno ghiacciato l’ aria, da lei non esce un sussulto, un segnale qualunque di vita. [AM]
 They-have frozen the air, from her not comes-out a shake, a sign QUALUNQUE of life
 ‘They have frozen the air, no shake comes out from her, no sign of life whatsoever.’
- (68) ...due corpi qualunque si attirano in modo proporzionale al prodotto delle loro masse
 ...two bodies QUALUNQUE themselves attract in way proportional to-the product of their masses
 ... [GEN]
 ‘...any two bodies attract each other in a way proportional to their masses ...’
- (69) Tiravo fuori un album, guardavo una o due fotografie al massimo, lo rimettevo a posto e ne prendevo un
 I-took out a album, I-looked one or two photos to-the maxim, it I-put to place and of-it I-took a
 altro, lo aprivo a una pagina qualunque e lo richiudevo subito. [IND-random]
 other, it I-opened at a page QUALUNQUE and it I-reclosed immediately.
 ‘I took an album, looked at no more than one or two photos, I put it away and I took another, I opened it at a random page and I closed it immediately.’

- (70) Quanti riguardi, ha urlato Antinoo, per uno straccione qualunque. [IND-average]
 How-many attentions, has screamed Antinoo, for a tramp QUALUNQUE.
 ‘How many attentions, Antinoo screamed, for just some tramp.’

Det + N + qualunque

- 7 tokens of *Det + N + qualunque* were found in the corpus, all of which exhibited the IND function.
- (71) Un’ antologia di quei film che hanno fatto di Sordi l’ immagine delle contraddizioni, delle debolezze, A antology of those films that have made of Sordi the image of-the contradictions, of-the weaknesses, delle virtù e della viltà dell’ italiano qualunque [IND-average]
 of-the virtues and of-the cowardice of-the italian qualunque
 ‘An antology of those films that made of Sordi the image of the contradictions, the weaknesses, the virtues and the cowardice of the average Italian’

Other

- One example was identified that could not be otherwise classified; it involves *qualunque* with the quantifier *ogni* ‘every’, and was classified as falling under the GEN function.
- (72) ...si legge: “Tirannide indistintamente appellar si debbe ogni qualunque governo in cui chi è ...one reads: “Tyranny indistinctly call one must every QUALUNQUE government in which who is preposto alla esecuzione delle leggi può farle, ... [GEN]
 in-charge to-the execution of-the laws can make-them, ...
 ‘...one reads: “Indistinctly one must call tyranny every government whatsoever in which the one who is in charge of the execution of the laws can also make them, ...’

5.5 Dutch

Data

- Indefinite forms included: *wh dan ook* series [*wat* ‘what’ / *welk(e)* ‘which’ / *wat voor* ‘what’ / *hoe* ‘how’ / *waar* ‘where’ / *wie* ‘who’ / *wanneer* ‘when’ / + *dan* ‘then’ + *ook* ‘also’]
- Corpus: Corpus Gesproken Nederlands CGN (Spoken Dutch Corpus; 10 million words; years 1998-2004)
- Query: *wat* [...] *dan* [...] *ook* / *welk(e)* / *wat voor* [...] *dan* [...] *ook* / *hoe* [...] *dan* [...] *ook* / *waar* [...] *dan* [...] *ook* / *wie* [...] *dan* [...] *ook* / *wanneer* [...] *dan* [...] *ook* (case insensitive). Here [...] indicates a gap of up to two words.
- Date of search: 11-5-2009
- Number of occurrences: 477 in total
- Labeled: All 477 occurrences loaded; 29 occurrences of *wie dan ook* annotated

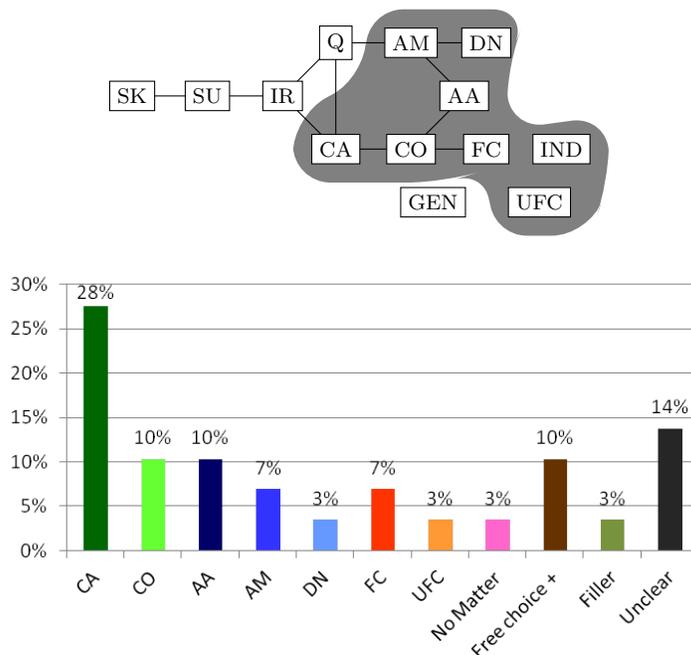
Annotation

- Item: *wh dan ook*
- Forms: *wat dan ook*, *welk(e) dan ook*, *wat voor dan ook*, *hoe dan ook*, *waar dan ook*, *wie dan ook*, *wanneer dan ook*
- Functions: Those described in Section 3 (including No Matter) plus FC+ (used for Free Choice with a restricted domain) and Filler (used for cases where *wie dan ook* serves as a filler)
- Syntactic tag: None used
- Semantic tag: None used
- Source: URL provided
- Date: Not available

Results

We report here on the annotation of *wie dan ook* in the Dutch synchronic data (29 tokens). Other forms of *wh dan ook* are included in the database but have not been annotated by function.

In present day spoken Dutch, *wie dan ook* has the distribution depicted below (note that the IND function is not found in the small sample of spoken data in our corpus, but is present in the written sample for the comparable period described in the diachronic section of this document, and thus for completeness is included in the graphic below):



For each of the functions for *wie dan ook* attested in present-day Dutch, a typical corpus example is given below. Note that some of these examples are taken from the written corpus data described in the diachronic section of this document; these examples are marked with the date of the original source:

- (73) Dat begreep de Amerikaanse legerleiding beter dan wie dan ook. (1995) [CO]
 that understand:PST.3SG the American army command better than WIE DAN OOK
 ‘The American army command understood that better than anyone.’
- (74) As de burgemeester of Pietje Patermans of wie dan ook mijn ’n mes in me knoke douwt en teuge me if the mayor or Pietje Patermans or WIE DAN OOK my a knife in me hands shove:3SG and to me
 zeit ‘Steek jij ouë Dries Valk ’s overhoop, dan haal ik z’n geeltjes ...’ (1920) [CA]
 say:3SG stab:IMP you old Dries Valk once to death then fetch:1SG I his banknotes
 ‘If the mayor or Pietje Patermans or anyone shoves me a knife in me hands and tells me: ‘Now you stab old Dries Valk to death, then I’ll get his money...’
- (75) Het is mijn werk, niet dat van de DG of van wie dan ook op het ministerie. [DN]
 it be:3SG my job not that of the DG or of WIE DAN OOK at the ministry
 ‘It is my job, not DG’s or anyone else’s at the ministry.’
- (76) Waar ik misschien wat al te scherp in mijn oordeel ben, wijte men dat niet aan lust om where I perhaps what all too sharp in my opinion be:1SG blame:3SG.SUBJ one that not on desire to
wie dan ook te kwetsen ... (1906) [AM]
 WIE DAN OOK to hurt:INF
 ‘Where I am perhaps a bit too sharp in my opinion, one should not blame that on a desire to hurt anyone ...’
- (77) Vele Edelen wijd en zijd erkenden en hingen hem aan; tegen wie dan ook de many aristocrats wide and side acknowledge:PST.3PL and hang:PST.3PL him on against WIE DAN OOK the
 Hoeksche partij haar geweld inrichtte. (1833) [AA]
 Hoeksche party her violence arrange:PST.3SG

'Many members of nobility, far and wide, acknowledged and supported him; against whomever the party 'Hoeksch' aimed its violence.'

- (78) Gesteld ik ging met een 100-tal schetsen, dat ik nu eens bijeen gezocht heb, naar postulate:PP I went with a hundredfold sketches that I now once together search:PP have:1SG to wie dan ook hier, op zijn best zou ik hooren vrees ik: 'dacht ge soms, dat die WIE DAN OOK here at his best will:PST.1SG I hear:INF fear:1SG I think:PST.2SG thee sometimes that those dingen handelswaarde hadden?' (1914) [FC]
things commercial value have:PST.3PL
'Suppose I went with about a hundred sketches, which I have put together for once, to anyone here, the best I fear I'd hear would be: 'did you think that those things have commercial value?'
- (79) Om slechts een voorbeeld te noemen, dat voor de hand ligt: wij vereeren Jan Steen en het nageslacht to just an example to name:I nftthat before the hand lie:3SG we admire:3PL Jan Steen and the offspring betaalt duizenden voor zijn schilderijen, en toch zou men zich de mogelijkheid kunnen voorstellen, pay:3SG thousands for his paintings and yet should:3SG one REFL the possibility can:IN imagine:INF dat Rembrandt of wie dan ook de richting van eerstgenoemde zou hebben veroordeeld en that Rembrandt or WIE DAN OOK the direction of the former will:PST.3SG have:INF condemn:PP and bestreden, al behoeft daarom onze hoogschatting van beiden in geen deele gewijzigd te dispute:PP although need:3SG consequently our esteem of both in none part change:PP to worden. (1883) [FC+]
become:INF
'To but name an example, that is quite obvious: we adore Jan Steen and future generations pay thousands for his paintings, and yet one could imagine the possibility that Rembrandt or whoever would have condemned or challenged the direction of the former, though that should not change our esteem for both.'
- (80) Koning Karel van Rumenie die een reformatie van het Hooger Onderwijs op het oog had en king Karel of Romania who a reformation of the higher education on the eye have:PST.3SG and incognito op een kasteel in de buurt logeerde – Falkenstein, meen ik – en die zoo verstandig was om incognito at a castle in the area reside:PST.3SG Falkenstein believe:SG I and who so wise was to bij wie dan ook, zijn licht op te steken, had mijn oom verzocht met mij, die hem vroeger at WIE DAN OOK his light up to put:INF have:PST.3SG my uncle request:PP with me who him in the past al eens in de zaak gesproken had, bij hem te komen. (1938) [UFC]
already once on the case speak:PP have:PST.3SG to him to come:INF
'King Karl of Romania, with his mind set on a reformation of higher education and incognito staying at a castle in nearby - Falkenstein, I believe - and as wise as to go to whoever/anyone/everyone for information, had asked my uncle to visit him with me, who had spoken to him before on this matter.'
- (81) Hoe kunt gij zoo dwalen? vraagt Mattheus, en hij verzekert u, niet dat wie dan ook den Heer uit how can:3SG thou so wander:inf ask:3SG Matthew and he assure:3SG you not that WIE DAN OOK the Lord from het graf heeft zien uitkomen, maar dat onder de oogen der romeinsche wachters een engel uit the grave have:3SG see:INF come:INF OUT but that under the eyes the:GEN Roman guards an angel from den hemel nedergedaald is, den steen van de grafdeur afgewenteld heeft, en is gaan zitten op the heaven descend:PP be:3SG the stone off the tomb's door roll:PP have:3SG and be:3SG go:INF sit:INF on dien steen. (1858) [IND]
that stone
'How can you be so astray? Matthew asks, and he assures you, not that just anyone has seen the Lord come out of the tomb, but that in front of the eyes of the Roman guards an angel has descended from heaven, has taken the stone off the tomb's door, and has taken his seat on that stone.'
- (82) En wie dan ook in de toekomst Zuid-Afrika zal regeeren, hetzij Boer of Brit, de twee rassen zullen and WIE DAN OOK in the future South-Africa will:3SG rule:INF either Boer or Brit the two races will:3PL samen moeten leven. (1902) [No matter]
together have:INF to live:INF
'And whoever will reign South-Africa in the future, be it Boer or Brit; the two races will have to live together.'
- (83) wel d'r is niet alleen wat de loutere inbreuken op het werk en wie dan ook met well there be:3SG not only what the mere infringements at the work and WIE DAN OOK with makelaars zonder erkenenis da's al een jaar gevangenisstraf ... [Filler]
real estate agents without acknowledgement that be:3SG already a year imprisonment
'Well there isn't only what the sheer infringements at work and whoever with brokers without acknowledgement that's already a year of jail ...

5.6 English

Data

- Indefinite items studied: *some*, *any*
- Corpus: British National Corpus (BNC) via BYU-BNC interface
- Query: any/some + singular noun
- Date of search: 18-1-2011
- Number of occurrences: 120505 occurrences of *any*; 165421 occurrences of *some*
- Number labeled: 100 instances of indefinites, including 80 instances of *any* and 20 instances of *some*

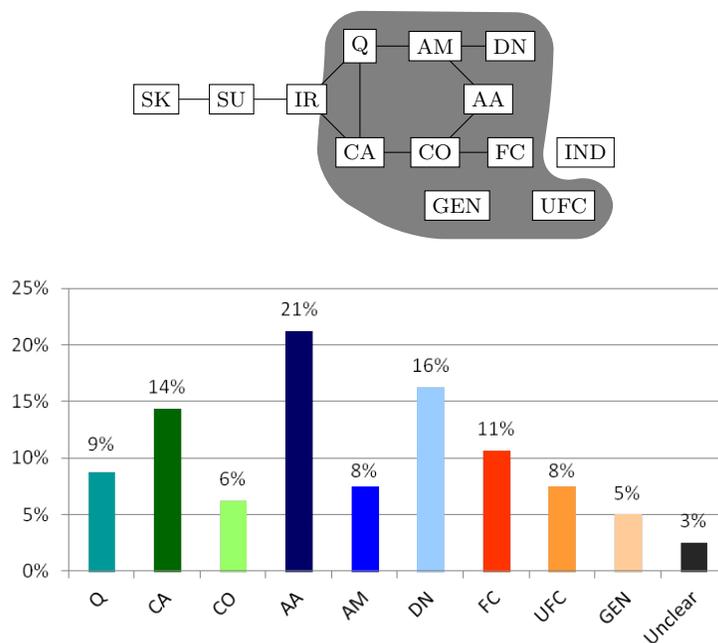
Annotation

- Items: some, any
- Form: none
- Functions: those described in Section 3 only
- Syntactic tag: None
- Semantic tag: None
- Source: included
- Year: included

Results

Here we described the results of the annotation of 80 *any* examples by a linguistically trained annotator. The results of a project in which 5 non-experts annotated all 100 **any/some** examples according to a set of prepared guidelines is described in Aloni et al. (2012).

As illustrated in our corpus data, English *any* has the distribution shown below:



While IND uses of *any* are attested in the literature (cf. Section 3 above), we did not identify any among the corpus tokens analyzed, likely due to the small sample size.

In the following, typical examples of each of function found in our data are shown:

- (84) Does a notion of “artistic growth” have any place in your scheme? [Q]
- (85) If you are in any way involved, potentially involved, or interested in ESPRIT CIME, you should make sure that PA are aware of you. [CA]
- (86) The reduction in child pedestrian fatalities may not therefore be quite what it seems, having perhaps more to do with parental sacrifice of the freedom of movement of their children than with any tangible gains in traffic safety. [CO]
- (87) The price of your holiday is fully guaranteed and will not be subject to any surcharges. [DN]
- (88) None of these references to smoking and cigarettes has any significance for the plot: they are paid advertising aimed mainly at children. [AM]
- (89) If these methods work, they may be used to modify fiduciary duties to avoid any conflicts with what is permitted or required by regulatory rules. [AA]
- (90) When you’ve mastered the various strokes possible (including lobs and smashes), you can enter any of the four Grand Slam tournaments (including Wimbledon), taking on computer players of varying ability. [FC]
- (91) In the social sciences at large the word has often been used very loosely for any approach which applies scientific method to human affairs, conceived of as part of the natural order. [UFC]
- (92) Sleep is a great help to any illness, so sleep as much as you can. [GEN]

5.7 Summary of synchronic results

The main prediction of Haspelmath were confirmed by our synchronic corpus research: there is no indefinite that violates the function contiguity. Further our study attested significant differences in meaning and distribution between the studied items which are often considered to belong to the same class of free choice indefinites in the linguistic literature.

6 Language-by-language data - diachronic

In addition to synchronic data, the *Corpus of Indefinite Uses* includes diachronic data tracing the development of indefinite forms in three languages: Spanish *cualquiera*, Dutch *wie dan ook*, and German *irgend*. The first two of these, in addition to the free choice meaning, share the property of being compounds and containing an interrogative word meaning ‘who’ or ‘which’ within their constituents. The third arose originally out of a particle with a locative meaning.

In this section, we describe the data available and their annotation, and summarize the central results of the diachronic analysis. For a more detailed discussion of these findings, see Aguilar-Guevara et al. (2011); de Vos (2010); Port (2012).

6.1 Spanish *cualquiera*

Data

- Indefinite form included: *cualquier(a)* [*cual-* ‘which’ + *quiera* ‘want:PRES.SUBJ.3’]
- Corpus: CORPUS DEL ESPAÑOL (by Mark Davies; 100 million words; spoken, fiction, newspaper and academic texts; 13th-20th centuries)
- Query: *ualq* [all possible forms of *cualquier(a)* + 10 instances of completely unrelated words, which were excluded]; search executed for the 13th century, 16th century, 18th century and 20th century, representing the four periods in which the history of Spanish has traditionally been divided (cf. Lapesa 1964). Note that the 20th century data included represent spoken language only (these are the data described in the synchronic section of this document)

- Date of search: 2009
- Number of occurrences: 13th century - 1012; 16th century - 5591; 18th century - 4048; 20th century - 7744
- Labeled: randomly chosen sample of 100 occurrences from each of the first 3 periods; 200 occurrences from the 20th century

Annotation

- Item: cualquiera
- Forms: none
- Functions: Those described in Section 3, including No matter (Spanish *cualquiera que sea*) + En cualquier caso ('in any case')
- Syntactic tag: None used
- Semantic tag: None used
- Source: Provided
- Date: Only century available; data annotated as 13, 16, 18, 20 (for 13th, 16th, 18th, 20th centuries, respectively)

Cualquiera (pronoun), or *cualquier* (determiner), translated to English as *whatever*, *whichever*, *whoever* or *any*, and composed of *cual* ('which/who') plus *quier(a)* ('want:3.PRES.SUBJ'), has been claimed to have emerged in Spanish as result of a grammaticalization process through which free relative clauses were reanalyzed as indefinite noun phrases (cf. Company-Company and Pozas-Loyo 2009).

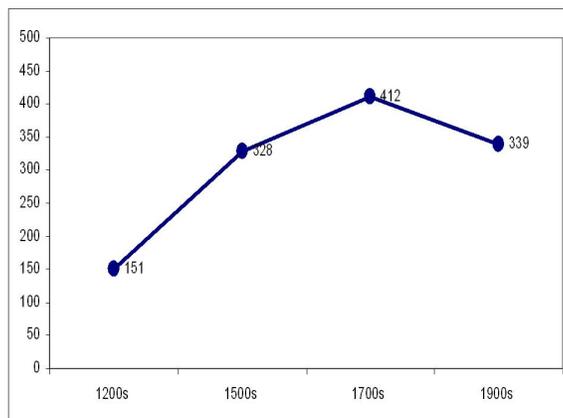
(93) Hypothesized grammaticalization process for *cualquiera*

- Free relative clause*
Haga en él **cual** castigo **quiera**.
do on him which punishment want:3.PRES.SUBJ
- Phrasal compound*
Haga en él **cual quiera** castigo.
do on him which want:3.PRES.SUBJ punishment
- Indefinite*
Haga en él **cualquier(a)** castigo
do on him whichever punishment

Presumably, this process occurred in the early stages of the history of Spanish and in consequence *cualquiera*, as a word, is already recurrently found in the first documentations of Spanish, which date back to the thirteenth century.

The following graph shows per million of words the number of occurrences of the construction in *El Corpus del Español* in the four periods selected for our corpus analysis.

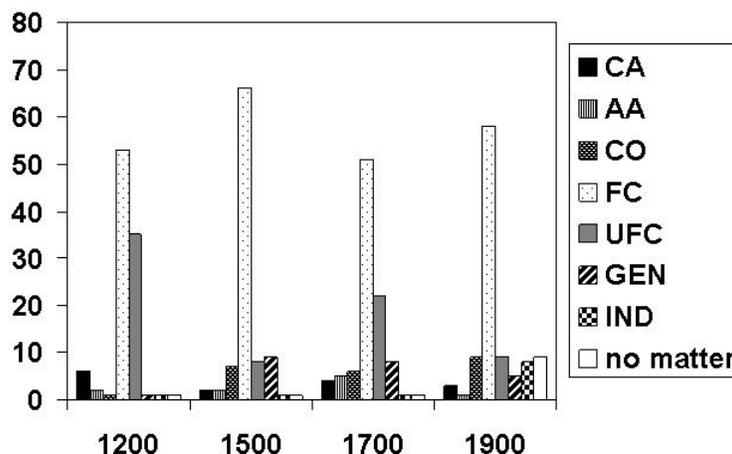
(94) Number of occurrences of 'cualquiera' per million of words.



As it can be seen, the frequency of *cualquiera* doubles between the 1200s and the 1500s, reaching a similar proportion to that documented for the 1900s. We take this as an indicator that the use of the construction has been consolidated at least since the sixteenth century.

The distribution of the functions that *cualquiera* covers throughout these periods points out to a similar conclusion:

(95) Functions covered by ‘cualquiera’ in 1200s, 1500s, 1700s and 1900s



The most noteworthy observation about this distribution is that, generally speaking, it has remained pretty similar throughout the four periods. The FC function is clearly the most dominant since the first period, with some other functions contiguous in the map (CA, CO and AA, UFC and GEN) having some presence as well. Interestingly, the UFC function displays a remarkable decrease starting in the 1500s. In Aguilar-Guevara et al. (2011), we tentatively attribute this to the fact that *cualquiera*, as part of its grammaticalization, occurs less and less frequently accompanied by post-nominal modifiers such as restrictive relative clauses and prepositional adjuncts, which typically serve as licenser of free choice items in UFC uses (e.g. *John kissed any woman #(with red hair)*). Finally, two more off-map functions, namely IND and *no-matter*, appear in the 1500s and gain presence by the 1900s.

Given the early grammaticalization of *cualquiera* and the stable distribution of its functions, our copus data does not provide much insight into the process by which this compound developed its current functions. This motivated us to study Dutch *wie dan ook*, which emerged more recently and which even today days appears to be less grammaticalized than *cualquiera*.

6.2 Dutch *wie dan ook*

Data

- Indefinite form included: *wie dan ook*
- Corpus: Written Dutch historical corpora - CD-ROM Middelnerlands (270 texts before 1300), DBNL (4458 texts from 1170-2010).
- Query: *wie dan ook* (case insensitive)
- Date of search: 21-04-2010
- Number of occurrences: 347
- Labeled: 347

Annotation

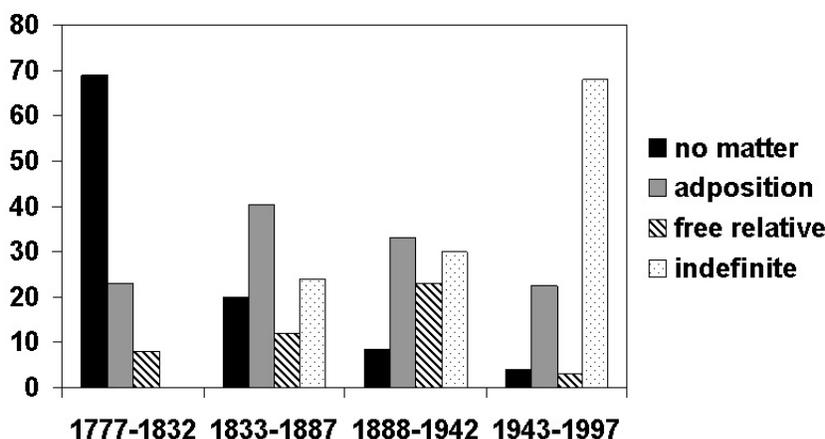
- Item: *wh dan ook*
- Forms: *wie dan ook*

- Functions: For coding of indefinites - those described in Section 3 plus ‘Free Choice +’ (used for Free Choice with a restricted domain) and ‘Filler’
- Syntactic tag: No matter, adposition, free relative, indefinite, unclear
- Semantic tag: To specify interpretation of No matter and Adposition examples - indifference, ignorance
- Source: Provided
- Date: Provided

Results

The first occurrence found for *wie dan ook* is from 1777; the period of this item’s existence has therefore been divided into four phases, each covering 55 years of the item’s evolution. The outcome shows that *wie dan ook* went through a four-stage process of grammaticalization:

(96) Four stages in grammaticalization of *wie dan ook*



Stage I The first phase in the grammaticalization of *wie dan ook* as an indefinite is formed by three forms of the *no matter*-function. Characteristic of types of *no matter* constructions is that *wh dan ook* is not part of the main clause yet: they all consist of either a *wh*-clause and a main clause, or a *wh*-clause within a main clause, as illustrated in the following:

- (97)
- Wie dan ook naar het feest komt; ik zal blij zijn.
‘Whoever comes to the party; I will be happy.’
 - [Wie dan ook naar het feest komt]_i; hij_i zal blij zijn.
‘[Whoever comes to the party]_i; he_i will be happy.’
 - Jan, (of) wie dan ook hij mag zijn, zal blij zijn.
‘John, (or) whoever he may be, will be happy.’

These forms occur around the same time. Together, they seem particularly frequent in the first phase, forming a significant majority of the total number of occurrences here, with this relative proportion decreasing in the three phases that follow (cf. the black bars in graph (96)).

Stage II In the following stage in the development of *wie dan ook* as an indefinite, *no matter*-constructions are shortened to adpositions, thus getting one step closer to becoming a grammaticalized indefinite. Adpositions have the following form: [..., [*wie dan ook*], ...]. They are shortenings of the *no matter*-function, formed by the ellipsis of the predicate. Although they do not form a separate *wh*-clause next to or within a main clause anymore, they are still not part of the actual sentence and therefore not real indefinites: they merely modify the noun they are placed after.

- (98) Als er iemand_i, wie dan ook_i, naar het feest komt, zal ik blij zijn.
 ‘If someone, whoever/anyone, comes to the party, I will be happy.’

As the grey bars in (96) show, this adpositional modification with *wie dan ook* (with ignorance or indifference meaning) is particularly frequent in the second phase in the development of this indefinite.

Stage III The third phase, the *free relative*-stage, shows a further integration of the *wie dan ook*-clause into the sentence, though still not a full integration either. The Free Relative (FR) function, the biggest part of the total number of occurrences of *wie dan ook* now, forms another spinoff of the *no matter* construction. However, whereas *no matter*-sentences still form combinations of wh-clauses (*wie dan ook* + predicate) and a main clause, the FR-function is more integrated than that, with the “*wie dan ook* + predicate” not forming a separate clause, but an actual part of the main clause, typically the subject. Examples of the FR-function have the following form: [[*wie dan ook* + predicate](,) VP], as illustrated in (99):

- (99) Wie dan ook naar het feest komt, zal blij zijn.
 ‘Whoever comes to the party(,) will be happy.’

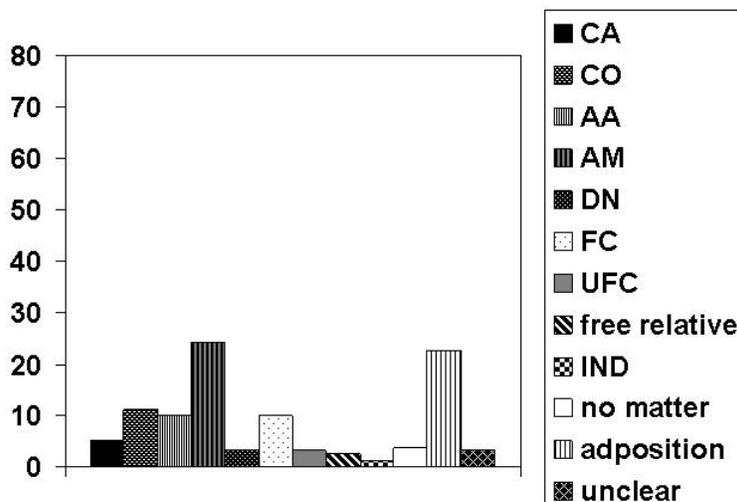
However, these subjects consisting of *wie dan ook* + predicate are often followed by a comma, thereby perhaps indicating that they are still seen as standing slightly outside of the actual sentence. Yet omitting the part starting with *wie dan ook* would give an incomplete thus ungrammatical sentence. This is a specific feature of the third phase; both the *no matter*-clauses and the adpositions can still be left out, of course sometimes causing a change in meaning of the sentence, but never with an incomplete sentence as a result. This shows how integrated a part of the sentence these occurrences of *wie dan ook* already form - although it apparently still feels a bit strange to the contemporary writer. Besides, these forms of *wie dan ook* are not as integrated yet as the plain indefinite will be.

Stage IV In this last stage of the grammaticalization of *wie dan ook*, the word group has finally become an indefinite. Examples of this kind form integrated parts of the sentence, with a plain *wie dan ook*, without any kind of predicate modifying it, being either subject or object: [... [*wie dan ook*] ...].

- (100) Je mag wie dan ook uitnodigen voor het feest.
 ‘You may invite anyone to the party.’

Indefinite uses of *wie dan ook* are attested from 1833 onwards, and their number increases in every phase, finally forming a vast majority of the occurrences in the fourth phase, as graph ((96)) illustrates. Here is the distribution of *wie dan ook* in stage IV:

- (101) Functions covered by ‘wie dan ook’ in stage IV



Summarizing: Overall, what can be concluded is that the process of grammaticalization of *wie dan ook* as an indefinite roughly followed four stages, starting off as a *no-matter* construction in a separate wh-clause, slowly evolving into an adpositional modifier on its own, while also turning into a part of the main clause with predicate, eventually yielding to the true and plain indefinite *wie dan ook* as part of a sentence.

6.3 German *irgend*

Data

- Indefinite form included: *irgend/irgendein*, in all declinations and spelling variants
- Corpus: Middle High German: Bochumer Middle High German Corpus (BC) - 102 texts/approx. 1,000,000 word forms; Middle High German Conceptual Database (MB) – 420 texts/8,598,568 words. Early New High German: Bonner Early New High German Corpus (BNHG), containing data from 1350-1700; supplemented by data from the text corpus of Thomas Gloning (<http://www.uni-giessen.de/gloning/etexte.htm>), Mediavum (<http://www.mediaevum.de/haupt2.htm>) and other electronic resources provided by wikisource.
- Query: BC, MB - *irgend* (lemma subsuming different spelling variants, e.g. *irne, jergendt, irgen, yrgend*). BNHG - **rg**. Supplemented by manual searches
- Date of search: 2009
- Number of occurrences: Middle High German - 109 (85 MB, 24 BC); Early New High German - 60 (BNHG 17, other sources 43)
- Labeled: All

Annotation

- Language Middle High German, EarlyNewHighGerman
- Item: *irgend**
- Forms: *irgend*
- Functions: Those described in Section 3; AA and AM functions collapsed under function IN (indirect negation)
- Syntactic tag: particle, intermediate, indefinite, unclear (to describe syntactic status of *irgend* form); superlative, restrictor (to indicate syntactic environment)
- Semantic tag: locative meaning, non-locative meaning, unclear, other (for particle syntax)
- Source: Provided
- Date: Provided

Results

Most of the diachronic literature on the history of German entertains a structuring into three or four periods, where the period of Middle High German is further divided into three substages, cf. Paul (2007).

three periods		four periods	
Old High German	750-1050	Old High German	750-1050
Middle High German	1050-1500	Middle High German	1050-1350
– Early Middle High German	1050-1170	– Early Middle High German	1050-1170
– Classical Middle High German	1170-1250	– Classical Middle High German	1170-1250
– Late Middle High German	1250-1500	– Late Middle High German	1250-1350
New High German	1500-present	Early New High German	1350-1650
		New High German	1650-present

In the corpus study, a four stage division was adopted. The results presented here are less fine grained and report only according to the two main stages : Middle High German and Early New High German.³

According to most scholars, the particle *irgend* emerged from the Old High German form *io-wergin*⁴ with a locative meaning similar to *anywhere/somewhere*. It developed later into a modifier of indefinites. Therefore, corpus examples were labeled with the status of *irgend*, being a particle or an indefinite. As an illustration, example (102) illustrates the particle use of *irgend*, while example (103) illustrates the indefinite use.

³A more detailed discussion of the results can be found in Port (2012).

⁴See for example Grimm and Grimm (1885), Lexer (1993), Krahe (1961). Fobbe (2004) assumes the form *io-hwar-gin* as the Old High German form. Irrespective of which form is assumed, all authors agree that the Old High German form for *irgend* had a locative meaning.

- particle:

(102) Die schult ich vf jne selbs lege, Tritt er *irgent* unszer dem wege.⁵
 the blame I on him self put, steps he *irgend* our the way.
 ‘I place the blame on himself, if he steps in our way somewhere’

- indefinite:

(103) und sollen sie deßwegen nicht zu Rede gestellet/ noch unter *irgend einen* Vorwand beschweret werden.⁶
 and should they therefore not to task take/ nor under *irgend a* pretext burdened be.
 ‘and they should therefore not be taken to task nor should they be burdened under any pretext’

Cases where it was not clear whether *irgend* was used as a particle or as an indefinite modifier were labeled as intermediate cases [IM]. The following example is a case in point:

- intermediate case [IM]:

(104) sahent ir *irgent* eynen ritter hie für ryten,⁷
 saw you *irgend* a knight here for ride,
 ‘did you see a knight **somewhere** here riding’ [particle]
 ‘did you see **some** knight riding here’ [indefinite]

As this example shows, the use as a particle with a locative meaning and the use as modifying the determiner *eynen* both lead to reasonable interpretations; thus the sentence is ambiguous.

Each token of particle *irgend* was further tagged according to whether it had a locative meaning or not. Notice that the data are annotated in a conservative fashion, which means that as long as a locative meaning for the particle was possible it was annotated as such. Example (102), repeated here as (105), illustrates a case in which a locative meaning for *irgend* is possible, though there might also be other readings available.

- particle, locative:

(105) Die schult ich vf jne selbs lege, Tritt er *irgent* unszer dem wege.
 the debt I on him self put, steps he *irgend* our the way.
 ‘The debt is on himself, if he steps in our way somewhere/ ever/ somehow’

A locative interpretation for the particle in (106) in contrast is impossible.

- particle, non-locative:

(106) ich sung im das allerpeste das ich *yergent* kan...⁸
 I sang him the very-best that I *irgend* could ...
 ‘I sang for him the very best I /ever/ somehow /# somewhere could ..’

Finally every sentence is annotated according to the original functional labels in Haspelmath’s implicational map, following an accordingly simplified version of the decision tree introduced in Section 3.

To summarize, there are 3 levels of annotation adopted here.

- Level 1: The status of *irgend*: particle vs. intermediate [IM] vs. indefinite
- Level 2: Meaning of the particle: locative vs. non-locative
- Level 3: Classification according to Haspelmath’s original functions.

At each level, examples that could not be classified were labelled as ‘unclear’.

⁵Diu Crone approx. 1325, MB

⁶Käyserl. und Frantzösischer Friedensschluß, 1679

⁷Prose-Lancelot, approx. , MB

⁸Neidhardt Lieder, app. 1210-1240, MB

Middle High German In most of the examples found for MHG *irgend* is used as a particle with a locative meaning. In the first example given here the context is the restriction of a universal (RL) and the reading can be paraphrased as *anywhere*.

- particle, locative, CA (RL):⁹

(107) da bekarte sich allez daz lut,, daz *irgen* in deme lande was,, zu unserme herren ihu xø(.)¹⁰
 Then converted *refl* all the people, that *irgend* at the country was to our Lord Jesus Christ.
 ‘At this time all the people who were anywhere in the country, converted to our Lord Jesus Christ.’

Example (108) shows an occurrence of the particle *irgend* in the antecedent of a conditional, again with a locative meaning:

- particle, locative, CA:

(108) ich biete dich durh dine gute., Swer mit dicheinen note *ierge* werde beuange., in
 ich ask-for you through your goodness., if-somebody with some distress *irgend* becomes frightened., in
 wage od’ in lande., also her miner helfe gere., daz ich in des muoze gewere. mit dineme troste.(.)¹¹
 waves or in land, as he my help desires., that I him it can grant. with your consolation.(.)
 ‘I ask you through your goodness, if anybody with any distress *anywhere* becomes frightened, on sea or on land, if
 he desires my help, that I can grant it with your consolation.’

A very frequent use of the locative readings of *irgend* shows up in combination with verbs of perception such as ‘see’ with a concrete object as an argument.

- particle, locative, Q:

(109) Hildebrant sach umbe sich, Ob er sîn hêrren Dieterich *Iergen* tôt ligen sæhe¹²
 Hildebrant saw around refl, whether he his master Dietrich *irgend* dead lying see
 ‘Hildebrant looked around, whether he would see his master Dietrich lying dead somewhere’

Another frequent construction is with superlatives, as illustrated by (110):

- particle, locative, CA (superlative)¹³:

(110) er ist ein recke uzerwelt und ist zu den besten gezelt, die *iergen* in dem lande sîn¹⁴
 he is a warrior chosen and is among the best ranked, who *irgend* in this country are
 ‘He is a chosen warrior and he ranks among the best who are anywhere in this country’

Now we turn to examples where it seems to be unclear what kind of reading the particle *irgend* exhibits:

- particle, locative, IR:

(111) Geliche buzze solle sie lyden,, wylge *irgo* gedar gan uzzet de clost’e,, ob it gedut
 Same penance should they suffer, who *irgend* attempted to-leave outside the nunnery,, when it happened
 ane der ebdissen gebot.(.)¹⁵
 without the abbess’s commandment.(.)
 ‘The same penance should they suffer, who *ever/somewhere* attempted to leave the nunnery, if it happened without
 the commandment of the abbess’

This sentence seems to favor a temporal reading rather than a locative reading. It states that nuns who leave the nunnery without the permission of the abbess should do penance. It seems not to be important *where* they might leave the nunnery, which would account for the locative reading, but if they *ever* do it without permission. But since

⁹The restriction of a universal are subsumed under the label CA, cf. the discussion in Port (2012).

¹⁰Frankfurter Predigtfragmente, end of 12th century, BC. All diachronic examples are given as found in the corpus, if not stated otherwise. That includes the spelling as well as the sentence or verse marker. In this specific example there are double commas which means that the annotator of this text confirms the comma which appeared in the original manuscript, the dot in brackets means that there is a dot missing in the original manuscript according to the annotator and the sentence ends there. Nothing hinges on this.

¹¹Trierer Aegidius, end of 12th century, BC

¹²Der Jüngere Sigenot, first quarter of 14th century, BC

¹³These superlative constructions are subsumed under the label CA, cf. discussion in Port (2012).

¹⁴Die Heidin, 1300-1325, MB

¹⁵Oxford Benediktinerregel, first half 14th century, BC

this sentence still makes sense with a locative meaning because the verb ‘to leave’ is involved which is compatible with locative modification, it was labeled as such. The next example instead clearly blocks a locative reading. Here the most natural reading seems to be a temporal one ‘*ever*’ or perhaps manner:

- particle, non-locative, CA (superlative):¹⁶

(112) ich sung im das allerpeste das ich *yergent* kan ...¹⁷
 I sang him the very-best that I *irgend* could ...
 ‘I sang for him the very best I *ever/somehow* could ...’

This example contains an individual level predicate, here being able to sing; such predicates do not readily combine with locative modifiers.¹⁸ Such a context seems to force the reader to find a suitable interpretation for the particle *irgend* different from the locative one.

Example (113) allows even more possible interpretations:

- particle, non-locative, CA:

(113) Han ich uch *irgent* erzúrnt, das sagent mir!¹⁹
 Have I you *irgend* enraged, that say to-me!
 ‘Have I enraged you *somehow/ some time /ever/#somewhere*, please tell me’

Here again the particle *irgent* seems to be best understood as a temporal or a manner modifier. Clearly the locative interpretation is dispreferred in this context.

In the next example the locative interpretation is clearly unsuitable. The most natural interpretation is the modal one. A temporal interpretation would not make sense either given the context.

(114) Und da er sich sahe so nacket, da schampt er sich vor imselber und bat sie das sie im
 And because he himself saw so naked, then feel-ashamed he *refl* of himself and asked they that they him
irgent geben ein cleyt, das er nit stuond also schemelich.²⁰
irgent give one piece of clothing, that he not stand too embarrassed
 ‘And because he saw himself being naked, he was embarrassed and asked, that they give him one piece of clothing
somehow, so ...’

Example (114) could be annotated as particle, non-locative, IR. But there is also the possibility that *irgent* modifies the DP *ein clyt* ‘*apparel*’ though *irgent* is not directly adjacent to the DP. If that is the case, sentence (114) could be paraphrased as follows in (115):

(115) Und da er sich sahe so nacket, da schampt er sich vor imselber und bat sie das sie im *irgent* geben ein cleyt, das er
 nit stuond also schemelich.
 ‘And because he saw himself being naked, he was embarrassed and asked, that they give him *some* piece of clothing,
 so ...’

In this case the label would be an indefinite, IR. As we have seen, this example allows both interpretations and therefore was labeled as an intermediate case [IM], more specifically as [IM], IR. These intermediate cases are ambiguous between modifying the verbal phrase (VP) or modifying a determiner phrase (DP).

The same ambiguity can be found in example (116), even if the particle is directly adjacent to the DP *ein ritter* ‘*one knight*’:

- intermediate [IM], CA:

(116) Der siebend sprach und gerett siner mynnen, ob yme *irgent* ein ritter begegnet der ...²¹
 the seventh spoke and promised his love, if him *irgend* a knight encounters who
 ‘The seventh spoke and promised his love, if him **irgend** one knight encounters who...’
 (i) ‘The seventh spoke and promised his love, if he encounters **somewhere/ ever** a knight who ...’ [particle]
 (ii) ‘The seventh spoke and promised his love, if he encounters **some** knight who ...’ [indefinite]

¹⁶The dots in this example are not in the original, they are mine showing abbreviation.

¹⁷Neidhardt *Lieder*, approx. 1210-1240, MB

¹⁸For the distinction between individual level predicates (ILP) and stage level predicates (SLP) and the interplay with different kind of modifications, cf. Kratzer (1995), Maienborn (2004), Vendler (1957).

¹⁹Prose *Lancelot* (part 1), app. 1250, MB

²⁰Prose *Lancelot* (part 3), app. 1250, MB

In the next example we find a prepositional phrase (PP) which is possibly modified by the particle. Again *irgend* is not directly adjacent to the PP and shows an intermediate stage [IM] with two possible readings:

- intermediate [IM], indirect negation [IN]:

- (117) " eya herre, durch gott, nit laßent mich sterben hie in diesem wald, sunder fürent mich irgent in ein
 " oh master, through God, not let me die here in this forest, but lead me irgent in a
 closter, da mir mög geschehen myn recht und das ich daselbst sterb als ein guote Cristen. ²²
 monastery, there me may happen my right and that I there die as a good Christian.
 'Oh master, through God, don't let me die in this forest, but lead me *irgend* in a monastery, there it may serves
 me right and I die there as a good Christian'
 (i) 'Oh master, through god, don't let me die in this forest, but lead me to *some* monastery, there I may receive
 what I deserve and I die there as a good Christian.' [indefinite]
 (ii) 'Oh master, through god, don't let me die in this forest, but lead me to a monastery *somewhere*, there I may
 receive what I deserve and I die there as a good Christian.' [particle]

Early New High German The first example given here from Early New High German clearly exhibits the particle reading we are already familiar with from the previous period of German. Since a locative reading is possible it was labeled as such.

- particle, locative, CA (superlative):

- (118) ...gegen dem Abent da wir bey Salamine in Hafen eingelauffen an welchem ort das beste Mohrsaltz so
 ...about the evening as we at Salamine at harbour arrived at which place the best salt so
irgend zu finden gemacht wirt²³
irgend to find produced is
 '... in the evening, when we arrived in the harbour of Salamine, where the best salt you find *anywhere* is produced'

Example (119) shows the same intermediate stage we saw in the phase of Middle High German: *irgend* could be understood as a particle and paraphrased as *somewhere* or it could be understood as modifying *ein beynelein* 'one bone'.

- IM, IR:

- (119) So nhu der tzan aus gebrochen ist sall man fleissigklich fuehlen ab sich auch *yrgent* ein beynelein von dem
 When *prt* the tooth out broken is should one carefully feel if *refl* also *irgend* a bone from the
 kinpacken abgescheldert hat...²⁴
 jaw splintered off
 (i) 'When the tooth is broken out one should carefully feel if *some* bit of bone has also splintered off from the jaw...'
 [indefinite]
 (ii) 'When the tooth is lost one should carefully feel if a bit of bone has also splintered off from the jaw *somewhere*...'
 [particle]

In the next example a locative interpretation for the particle makes less sense, but still it would not be impossible. The most natural reading seems to be that *irgend* modifies *ein Schmerz*, but due to the conservative annotation, this example was still annotated as an intermediate case:

- IM, Q:

- (120) Ach? bedenket doch nur/ ob *irgend ein* Schmerz sei/ als mein Schmerz/ und ein Jammer/ als der Jammer/
 Oh? consider *prt prt*/ if *irgend* a pain is/ like my pain/ and a sorrow/ as the sorrow/
 der mich treffen hat.²⁵
 which me hit has
 (i) 'Oh, consider, if there is anywhere a pain like my pain, and a sorrow as the sorrow which has hit me' [particle]
 (ii) 'Oh, consider, if there is *irgend-one* pain like my pain, and a sorrow as the sorrow which has hit me' [indefinite]

²¹Prose Lancelot (part 2), app. 1250, MB

²²Prose Lancelot (part 3), app.1250, MB

²³Leonhart Rauwolf: Beschreibung, Lauingen 1582

²⁴Artzney Buchlein wider allerlei krankkeyten vnd gebrechen der tzeen, 1530

The next example clearly shows an indefinite use:

- indefinite, Q:

- (121) ...vnd sprachen zu jm / Jsts auch recht / Das sich ein Man scheidet von seinem Weibe / vmb *irgend eine*
 ...and spoke to him / is also right / that refl. a man divorced from his wife / for *irgend-one*
*ursache?*²⁶
 reason?
 '...and ask him/ Is it right/ that a man gets divorced from his wife/ for *any* reason?'

In general, we find the same intermediate stage as in Middle High German, extending now to other indefinite expression like *etwas* 'something' or with the indefinite pronoun *einer* 'someone'. The following example is a partitive construction with the pronoun *einer* 'someone':

- indefinite, indirect negation (IN):

- (122) Drauff jhm Don Quixote antwortet / daß er nicht ein Scherff bey sich hette / vnnd daß darumb / weil er
 then him Don Quixote answers / that he not a red-cent with him has / and that because / because he
 niemals in einiger Geschichtbeschreibung der wallenden Ritter gelesen / daß *irgend* der Ritter *einer* Geld
 never in a history the roaming knights read / that *irgend* the knights ones money
 bey sich geführt hette.²⁷
 with *refl* have had.
 'and that because he has never read in any history of the roaming knights, that *anyone* of the knights had money
 with him.'

In the following example *was* 'what' is modified by *irgend*:

- indefinite, IN?:

- (123) Viel können schwerlich leiden, daß von anderen Leutten auch *irgend was* löbliches geredet oder gerühmet
 Many can hardly bear, that of other people also *irgend was* worthy said or praised
 werde, es verdreißt sie solches im Hertenzen, als ob ihnen etwas von ihren Ehren damit benommen
 is, it mortifies them so in heart, as if them something of their honesty thereby withdraw
 wäre.²⁸
 would.
 'Many can hardly bear, that something good is also said about other people or praised, it mortifies them so in their
 heart, as if something of their honor would be taken away by it.'

The next example shows that the use of *irgend* is also extended to the *wh*-pronoun *where* in Early New High German.

- indefinite, CO:

- (124) Auff der kleinen Seyten zu Prag wird so gut Teutsch geredet / als irgendswo in gantz Teutschland; das
 At the small sites to Prague is as good German spoken / as irgendswo in whole Germany: that
 macht / daß die Teutschredende keine baurische Nachbarn auff den umbligenden Dörffern haben / die ihnen
 is / that the German-speakers no rustic neighbors at the close villages have / who them
 ihre Sprach verderben; dahingegen die Franckfurter von den Wetterauern: die Straßburger von den
 their language ruin; contrary the Frankfurter by the Wetteraurn: the Straßburger by the
 Kocherspergern:...²⁹
 Kochersbergern: ...
 'In the small villages close to Prague one speaks the same good German as *anywhere* in the whole of Germany:
 that is because the German speakers have no rustic neighbors in the small villages close by who could ruin their
 language; in contrast the Frankfurter can be ruined by the Wetteraurn: the...'

The composition of *irgend* with *wo* shows that the particle *irgend* is not necessarily understood as having a locative meaning anymore, otherwise combining these two items would be redundant.

Example (125) could be an early case of *irgendein* in the specific-unknown function.

²⁵Georg Göz: Leich-Abdankungen, Jena 1664

²⁶Luther letzte Hand, 1545. This is one of the 5 examples for *irgend* Fobbe (2004) represents in her diachronic research.

²⁷Miguel de Cervantes : Don Kichote de la Mantzscha, Franckfurt, 1648

²⁸Hans Michael Moscherosch: Gesichte, Straßburg 1650

²⁹Hans Jakob Christoffel von Grimmelshausen: Deß Weltberuffenen Simplicissimi Pralerey und Geprßng mit seinem Teutschen Michel, 1673

- indefinite, SU:

(125) ... also wurden wir getrieben in die Flache deß Meerbusens bey Guanipa, von dannen den Mundt
 ... then were we driven in the shallows of the delta of gulf by Guanipa, from there into the mouth
irgendt eines Flusses die Johan Dowglaß zuvor hatte erkundigt (zu erreychen). Wir hatten auch einen
 irgend one river which Johan Dowglaß before had explored (to reach). We had also an
 Indianischen Pilot bey vns...³⁰
 Indian pilot with us...
 ‘then we were driven into the shallows of the delta by Guanipa, from there into the mouth of *some* river which
 Johan Dowglaß explored before (to reach). We had also an Indian pilot with us...’

The example reports on a travel in America. It seems that the writer reports that there is a river they reached which was previously explored by Johan Dowglaß. This would be compatible with a specific unknown reading (SU).

Discussion In the stage of MHG, *irgend* was mainly used as a particle with a locative meaning. At the same time, the first intermediate cases appeared, but no cases in which the particle modifies an indefinite phrase unambiguously. The first occurrences for such uses were found in the stage of Early New High German. In this stage, *irgend* does not only modify indefinite phrases based on the determiner *ein*, but is also extended to the pronoun *einer* and *wh*-based indefinite phrases.

The following table shows the distribution of the particle, the intermediate uses [IM] and indefinite according to Level 1 in the annotation procedure:

(126) *irgend and its status in Middle High German and Early New High German*

period	total	particle	IM	indefinite	unclear
MHG	109	90 (82,6 %)	17 (15,6 %)	0	2 (1,8 %)
ENHG	61	25 (41 %)	9 (14,7 %)	27 (44,3 %)	0

At Level 2 the 86 occurrences of the particle were annotated w.r.t. their locative or non-locative meaning in the context. The following graph displays the results of this labeling. For MHG, 8 uses of the particle are unclear. For ENHG, there were three cases of the particle used in an approximate phrase or having the meaning of *perhaps* and are subsumed under the label *other* in the corpus. Four cases are unclear.

(127) *particle: locative vs. non-locative readings in MHG and ENHG*

period	particle	locative	non-locative
MHG	90	66 (73 %)	16 (17,7 %)
ENHG	25	13 (52 %)	5 (20 %)

The following tables show the distribution of the particle according to Haspelmath’s functions.

(128) *Distribution of the particle in Middle High German*

functions	occurrences	percentage
IR	22	20,2 %
Q	22	20,2 %
CA	54	49,5 %
IN	8	7,3 %
UN	3	2,8 %
in total	109	100 %

The intermediate cases [IM] cover the same area on the implicational map as the particle in MHG. The following table shows only the distribution of the *irgend*-indefinites in Early New High German.

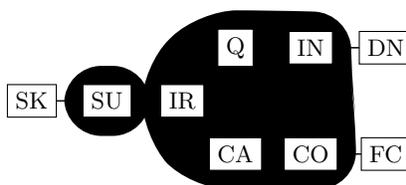
³⁰Walter Raleigh: Amerika, Frankfurt/Main 1599

(129) *Distribution of the irgend-indefinites in ENHG*

functions	occurrences	percentage
IR	8	29,6 %
CA	5	18,6 %
IN	8	29,6 %
Q	1	3,7 %
CO	4	14,8 %
SU	1	3,7 %
total	27	100 %

As can be seen from table (129), the *irgend*-indefinites show up in a wide area of the implicational map.

(130) *distribution of irgendein in Early New High German*



References

- Ana Aguilar-Guevara, Maria Aloni, Angelika Port, , Radek Šimík, Machteld de Vos, and Hedde Zeijlstra. *Semantics and pragmatics of indefinites: methodology for a synchronic and diachronic corpus study*, 2011.
- Maria Aloni, Andreas van Cranenburgh, Raquel Fernandez, and Marta Sznajder. *Building a corpus of indefinite uses annotated with fine-grained semantic functions*. In *Proceedings of LREC 2012*, 2012.
- Luis Alonso-Ovalle and Paula Menéndez-Benito. *Modal indefinites*. *Natural Language Semantics*, 18:1–31, 2010.
- Concepción Company-Company and Julia Pozas-Loyo. *Los indefinidos compuestos y los pronombres genérico-impersonales omne y uno*. In Concepción Company-Company, editor, *Sintaxis histórica de le lengua española (Segunda parte: La frase nominal)*, pages 1073–1219. Fondo de Cultura Económica-Universidad Nacional Autónoma de México, México City, 2009.
- Veneeta Dayal. *Any as inherently modal*. *Linguistics and Philosophy*, 21:433–476, 1998.
- Machteld de Vos. *Wh dan ook: The synchronic and diachronic study of the grammaticalization of a Dutch indefinite*. BA thesis, University of Amsterdam, 2010.
- Eilika Fobbe. *Die Indefinitpronomina des Deutschen. Aspekte ihrer Verwendung und ihrer historischen Entwicklung*. Universitätsverlag Winter, Heidelberg, 2004.
- Anastasia Giannakidou. *The meaning of free choice*. *Linguistics and Philosophy*, 24:659–735, 2001.
- H. Paul Grice. *Logic and conversation*. In P. Cole and J. L. Morgan, editors, *Syntax and semantics, Vol 3: Speech acts*, pages 41–58. Seminar Press, New York, 1975.
- Jacob Grimm and Wilhelm Grimm. *Deutsches Wörterbuch*. 1885.
- Martin Haspelmath. *Indefinite pronouns*. Oxford University Press, Oxford, 1997.
- Laurence Horn. *Airport '86 revisited: Toward a unified indefinite any*. In G. Carlson and F.J. Pelletier, editors, *The Partee Effect*, pages 179–205. CSLI, Stanford, 2005.

- Jacques Jayez and Lucia Tovena. Epistemic determiners. *Journal of Semantics*, 23:217–250, 2006.
- Dr. Hans Krahe. *Germanische Sprachwissenschaft. Formenlehre*. Walter de Gruyter & CO., Berlin, 1961.
- Angelika Kratzer. Stage Level and Individual Level Predicates. In Gregory Carlson and Francis Jeffrey Pelletier, editors, *The Generic Book*, pages 125–175. The University of Chicago Press, Chicago, 1995.
- Angelika Kratzer and Junko Shimoyama. Indeterminate pronouns: The view from Japanese. In Yukio Otsu, editor, *The proceedings of the Third Tokyo Conference on Psycholinguistics*, pages 1–25, Tokyo, 2002.
- Rafael Lapesa. *Historia de la lengua española*. Gredos, Madrid, 1964.
- Matthias Lexer. *Mittelhochdeutsches Handwörterbuch*. S. Hirzel, Stuttgart, 38 edition, 1993.
- Claudia Maienborn. A Pragmatic Explanation of the Stage Level/Individual Level Contrast in Combination with Locatives. In Brian Agbayani, Vida Samiian, and Benjamin V. Tucker, editors, *Proceedings of the Western Conference on Linguistics (WECOL) 15*, volume 4, Fresno: CSU, 2004.
- Paula Menéndez-Benito. On universal Free Choice items. *Natural Language Semantics*, 18:33–64, 2010.
- Hermann Paul. *Mittelhochdeutsche Grammatik*. Max Niemeyer Verlag, Tübingen, 25th edition, 2007.
- Angelika Port. The diachronic development of German *irgend-indefinites*. Ms, 2012.
- Zeno Vendler. Verbs and Times. *The Philosophical Review*, 66(2):143–160, 1957.
- Frans Zwarts. Three types of polarity. In F. Hamm and E. Hinrichs, editors, *Plurality and quantification*, pages 177–238. Kluwer, Dordrecht, 1998.