SSWL and Terraling: some supporting documents

All property definitions on the Terraling site site give explicit instructions, elicitation contexts and examples.

This document does not include additional supporting materials for:
navigating Terraling (navigating Terraling

Subject properties: pivots/subjects in existential constructions here and(
here
definite and generic subjects (supporting documents in progress)
properties in current development (Ivory Coast)

Cristina Guardiano
and
Hilda Koopman

July 10, 2018

Contents

1 Why SSWL/ Terraling? An introduction.

1.1 Need for a syntactic/semantic database that can support current and future theoretical research

- Internal to formal linguistics

  - The role of comparative syntax and future theoretical development.
  - Possible application of current available and future research tools (phylogenetics,...)

- External to formal linguistics

  - Database to complement and offer an alternative to WALS.
  - Database were variability is not hidden, properties are finegrained, which do not only code what we already know, but allow future discovery and exploration.
  - General education of the public at large.

SSWL: (Syntactic (and Semantic) Structures of the World's Languages), (Chris Collins, Richard Kayne), Hilda Koopman and Dennis Shasha (Prof Computer Science NYU, plant genomics)). Pilot: Pilot: http://sswl.railsplayground.net/ Current: Terraling http://test.terraling.com/ **Goal**: develop an open-ended

database (for the community by the community) of the (morpho-)syntactic structures of the world's languages that can serve as a tool in support of fundamental research. *Vision: build a "genomics" database of the (morpho-) syntactic/semantic properties of human languages.* This project is directly related to strongly decompositional approaches to syntax, and cartography, "maps of structures".

Provide over time as detailed cross-linguistically comparable data as needed:

- to document the internal structure of such systems
- to document all the basic ingredients needed for their description
- to record the extent and limit of their variation

Terraling– a collection of datasets. http://test.terraling.com SSWL Hilda Koopman (UCLA linguistics), Dennis Shasha (NYU), Hannan But (back-end), Marco Liberati (back-end), Connor Mayer (UCLA linguistics, Tyler Carson (UCLA linguistics)

Data base functionality:
Building a database in support of theoretical research is a longtime project: the database must be able to last, allow for new content and new languages to be added, which, in the case of SSWL must involve linguistic experts (expert crowdsourcing).

- store the data in a secure world-wide accessible database-backed website
- allow flexible additions to data as new properties and new languages are added

- allow disciplined and secure curation of data by multiple linguists (user friendly, for contributors, administration, and the general research community)
- allow sophisticated queries.

Relational database with highly flexible database scheme.

(1)   Language, property, value (yes,no)
        Italian, SV, yes
        Italian, VS, yes

2   Current data sets

- Database schema can be adapted to anything, as there is no preconceived idea of what Language is.
- This means it can be adapted to other linguistic (as well as non linguistic) research projects, depending on the research questions one asks
- Or code up data from other datasets (all data from SAND have been previously translated into the Terraling dataset format.

Current datasets (which can have a private setting (for the duration of a project or dataentry) or a public setting):

http://test.terraling.com/3
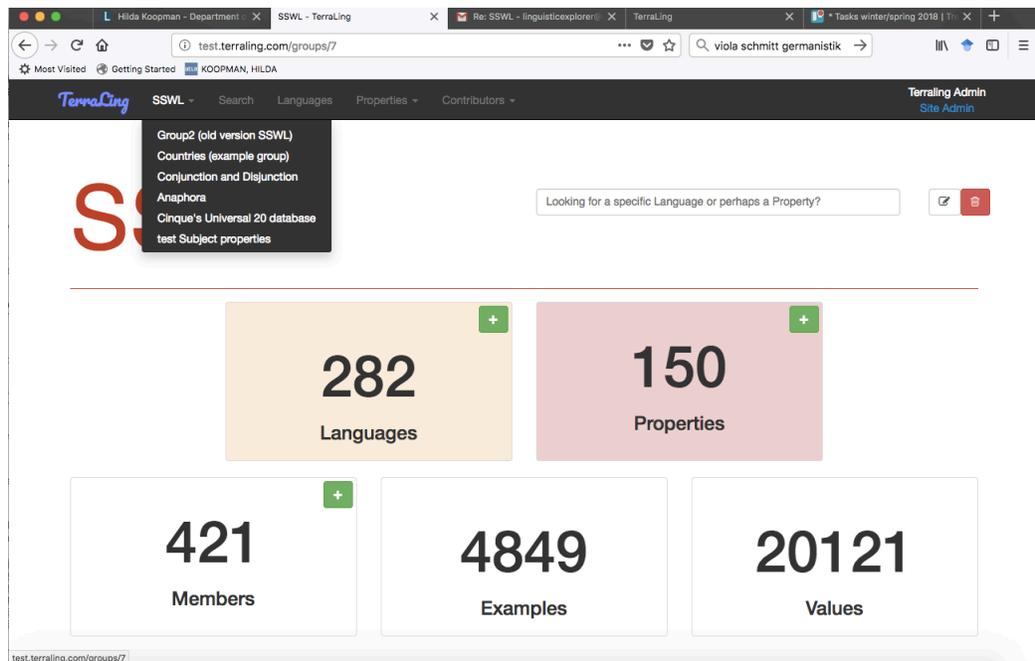
- SSWL (we are ready to post approximately 40 subject properties (form (bare, article, etc)- meaning (definite, indefinite, generic)- type of noun (mass, count sg/pl,)) properties in mid July 2018).

- *Conjunction and disjunction from a typological perspective* (semantics) PI: Viola Schmitt (University of Vienna).http://www.univie.ac.at/konjunktion/texts.html

- Anaphora http://grantome.com/grant/NSF/BCS-1424336.

- Cinque's Universal-20 database

- Several individual projects are in development...

3 Structure of current SSWL properties (as of July 10, 2018.)

The SSWL database is an ever-expanding database: new properties can be added. Important to get an idea of what the 150 (soon 190) current properties in the SSWL database are.

The data range from very broad typological properties concerning word order *..Explain how these are different from Wals* http://wals.info/chapter/s6 to queries about quite detailed properties of (some) regions that make up the noun phrase.

3.1 (some) Basic word order typology: 32 properties

Properties come in groups that are thematically related, and ask for binary values (yes/no (or NA (non applicable), as described in the property definitions. (SSWL codes the occurrence of particular orders in a language, given a specific definition, and does not code for "dominant" order. Discussion on the notion of domination order Basic properties include so far (a subset of) basic word order properties and

some properties about agreement, familiar from the typological literature, starting with Greenberg (1965).

3.2 How is SSWL different from WALS here? On the notion of "dominant" order and problems related to "dominant"order.

This subsection explains a crucial difference with WALS http://wals.info/feature/87A#2/18.0/152.9. in SSWL, a yes value obtains when a particular order is attested (as defined in in the query).

An example: 13_A N Definition: The property "A N" (Adjective Noun) has the value "yes" when an adjective can precede the noun it modifies in a neutral context. This definition concerns attributive adjectives that modify a noun such as "big ball", not predicate adjectives such as "big" in the sentence "The ball is big." The noun should be a common noun such as "book", "person", "house", etc. (as opposed to a proper noun, pronoun, or quantificational expressions like "something"). As with all word order properties, we restrict our attention to productive word order patterns.

General Note on Word Order Properties: A "yes" value for "A N" Adjective Noun does not exclude a "yes" value for Noun Adjective. In other words, a language could allow both orders in a neutral context.

How would we code French?

(2)   A(djective) Nom and Nom A(djective).

a.   une jolie fleur
b.   une jolie fleur rouge

	N A	A N	
SSWL	yes	yes	French
WALS	yes	no	French

The SSWL properties do not "bin" orders: i.e. they simply record what orders are attested given a particular definition: Important: the notion of **dominant order** does not play any role in SSWL.

What is an Adjective?
Comment définir des catégories? (cf atéliers)

NB: The original properties asked to coded languages in which adjectives look like relative clauses as N(ot)A(pplicable).

Possible orders for Adjective Noun orders dans la base de données:

Prop	Value	Prop	Value	Count
13_AN	No	14_NA	Yes	?
13_AN	Yes	14_NA	Yes	
13_AN	Yes	14_NA	No	
13_AN	NA	14_NA	NA	
13_AN	No	12_NA	No	

What is the current distribution of values in the database?
We can answer this question by going to Search on the Masthead.
Select properties 13 and 14,
click on Cross (Cross can extract all data from up to 6 properties)
The results of Cross can be downloaded, Saved, and accessed through History, which will update the search for the state of the database. Terraling cross-search (SSWL data, 7-25-2017)

Prop	Value	Prop	Value	Count
13_AN	No	14_NA	Yes	90
13_AN	Yes	14_NA	Yes	80
13_AN	Yes	14_NA	No	57
13_AN	NA	14_NA	NA	3
13_AN	No	12_NA	No	0

- Note that this is not a balanced language sample, it simply reflects what is found for the languages in the database.

- This will allow micro-comparisons between closely related languages, ei- ther historically or synchronically. It is possible to do the same search for a balanced sample.

- This table of course says nothing about the analysis of the data!
  For a spelled out theories which generates testable hypotheses, see Koopman Adjectival questionnaire). 3 languages have the property NA (non applicable). (Warlpiri, Chicasaw

6

3.3   Negative properties

There is a block of negative properties, concerning order, and form.(submitted by 15 minutes)

3.4   Polar questions and answers

There is a set of properties for polar questions, and answers (written by Anders Holmberg and Craig Sailor) (25 minutes)

3.5   V-second and Aux selection

(10 minutes) V second: pas relevant pour les langues africaines

3.6   Universal 20 properties (Dem A N and Dem Num A)

Order of Dem Num A N
So far: Greenberg (1965), Cinque (2005).

- simple combinations: AN, and NA; Dem N and N Dem, Num N and N Num.
- combinations of 3 properties: 2 sets of 6 (12 properties) (10 minutes)
- A set of 4 properties for indefinite and definite Num N combinations.
  There are three children and two cats on the street.
  The three children are playing soccer.
  The two cats are watching the children. ...with follow up properties about form (initial attempt for determiner system and definiteness and indefiniteness in the noun phrase): (10 minutes)

3.7   Properties of definites and indefinite noun phrases

(see next chapter)

3.8   Properties in various stages of development

- demonstratives (cristina Guardiano)
- definiteness marked outside of the noun phrase/clitic doubling?
- citic combinations (PCC)
- scope of indefinites w.r.t. negation
- (position) subordinators
- verbal possession– and the verb need (Stephanie Harves)
- compounds
- on culmination/non-culmination accomplishments (Lisa Travis)
- relative clauses
- adjectives (adjectival questionnaire) and something interesting
- "other"
- ordinals
- superlatives
- causatives...

4 Towards a cartography of the D region: Bare nouns and "Articles" (Demonstratives, case, classifiers, and noun classes.)

Cristina Guardiano and Hilda Koopman. Please Read this document before you start the object properties.

This short document discusses three aspects of the documentation project which aims to code certain properties of the D region, related to the distribution of **bare** nouns and **determiners/articles**, and their interpretation. This is an area known to exhibit much crosslinguistic variation, and of considerable interest to the syntax/semantics literature.

This requires defining what we understand by *article*, and what we understand by *bare nouns*, which we do below.

So far, we posted:

- a set of 40 object properties (subdivided in groups of 3, 4, or 5, as explained below) have been posted that probe for a particular set of phenomena related to nominal structures in object position.

- 37 properties for subject position (including properties of existential constructions, and negative existential constructions)

- 16 properties related to vocatives.

Usually, a group of property can be set on the basis of quite simple elicitation task, that asks for examples given a particular scenario/context, and then asks how to classify the examples according to a number of variables.

4.1 Variables included in the properties

The properties are organized around the following variables: which can be further expanded, or refined

1. indefinite vs. definite reading of (unmodified) noun phrases, and generic readings for:

   (a) mass nouns vs. sg/pl count nouns
   (b) nouns with (intrinsically) unique reference (*sun, sky, moon*)
   (c) proper names, proper names modified by adjectives

   with respect to syntactic position:

   (a) syntactic position: object
   (b) syntactic position: object
   (c) vocatives

2. ordering properties (if there is an "article" does it precede the N or follow?)

3. subject properties (36 properties)

- defart-N, N defart, art –proper name, proper name–article etc...
  to complete, still missing: Order indefinite articles-N, and scopal properties for indefinites
- order Modifier-Proper Names

In particular, we are interested in whether a nominal expression, according to each of the variables 1-6/ 1-5-7:

    a.   can be "bare" (i.e. lack an "article"), or
    b.   must have an "article", or
    c.   can have an article (i.e. can either be bare or occur with an article)

This gives rise to different property definitions for objects, organized in the following fashion. The task (defined below) asks for yes/no values:

(3)   Indefinite mass nouns in object position

    a.   can be bare   yes/no
    b.   must have an article   yes/no
    c.   can have an article   yes/no

(4)   definite mass nouns in object position.

    a.   can be bare   yes/no
    b.   must have an article   yes/no
    c.   can have an article   yes/no

and so forth.

These properties apply to:

- unmodified mass/count nouns (NO quantifiers, adjectives, possessors, relative clauses, adpositional complements)
- objects in affirmative transitive sentences (NO negative, interrogative, passive)

Each **property definition**: is a small field-work questionnaire with instructions on how to set the value, elicitation contexts and illustrations, and is organized in the following way.

- definition of the property
- elicitation context(s)
- explicit instructions on how to set the values for the language in question
- examples that illustrate the yes/no values for different languages
- comments

Definitions: What counts as bare and what counts as an "article"? This is necessary because of the absence of stable linguistic definitions: the current terminology depends on the language, but does not allow crosslinguistic comparisons For the purposes of this set of properties, we do this as follows (this is repeated and slightly adjusted in each property definition),

A bare noun phrase ....

- for languages with definite/indefinite articles, specificity markers, definite/indefinite affixes:
  lacks an article (can be bare)
- for languages without articles but with case, adpositional affixes/endings, classifiers, noun classes/augments,
  exhibits no structural alternation with noun phrases of a particular type (defined in each property definition)

An "article" .....

- self explanatory for languages with articles
- for languages without articles: sometimes "case", adpositions, count as "article" (if case expresses a contrast definite and indefinite readings of objects.)
- Demonstratives counts as an "article" if they can be used like articles (and should show homogeneity (aka as consistency) effects (Löbner, 2000))
- Classifiers can count as articles (in certain readings/circumstances).
- noun class markers/ augments may count as articles in a subset of contexts.

The way to proceed: read the property definitions.. (these define what you are supposed to do)

(5)   a.   step 1: "fieldwork": elicit examples in language X. (Questionnaire, with elicitation sentences and contexts;

        b.   step 2: construct the examples (can be more than one per property), paying attention to the property definitions.
Type up the examples in the format given, (and feel free to send them to me and Cristina for comments).

        c.   step 3: Classify the examples: read property definitions carefully, and set the property values. (Use the supporting excel sheets..(we still need to update these))
Can the task be done in your language? note any questions/ unclarities, and run them by us.

        d.   step 4: (creating the language and) enter the property values, with examples, and comments into the database (we added a *how sure are you* measure in Terraling.
Comments about distribution are particularly important. In many cases further work will be needed, please feel free to indicate so !

        e.   step 5: explore with the database search tools. (Searches include: complex searches, implications, compare (up to 8 languages), similarity trees, map results, download, saved searches with stable handle (when one returns to the database the searches will update...)

Findings and Theory

The property definitions define 3 types of languages for each set of a/b/c/:
3 types of languages are excluded.

(In)def sg Count N	can be bare	must have article	can have article	found?
	yes	no	yes	found
	no	yes	no	
	yes	no	no	
*	yes	yes	no	0 *contradictory*
*	no	no	yes	0 *contradictory*
*	no	yes	yes	0 *contradictory, for same article*

Note that contradictory statements hold only for the " same" article. Since we classify some expressions of case as an article, and similarly we classify some uses of Demonstratives as articles, we allow for languages that have an obligatory case marker for objects expressing definiteness, and an optional Dem yo yield both " must" Yes and " Can have"

4.2   Results and discussion

Write to me if you want the draft.

5 Negative existential contexts: on the form(s) of indefinite subjects: bare, articles, NPIs, "special" forms.

Cristina Guardiano[1]
and Hilda Koopman[2]
Version of July 10, 2018. (This document lacks some references.)

Each property definition includes instructions, an elicitation contexts, examples of values in individual languages. A file with examples to elicit can be found here. PLEASE READ THIS DOCUMENT TO GET AN IDEA OF THE GENERAL CODING SCHEMA BEFORE CODING THE NEGATIVE EXISTENTIAL PROPERTIES.

1. WE PRESUPPOSE FAMILIARITY WIH THE CODING SCHEMA FOR OBJECT PROPERTIES here , WHICH INCLUDE OUR DEFINITION WHAT COUNTS AS AN ARTICLE, AND EXISTENTIAL CONSTRUCTIONS here.

5.1 Background

Here we are interested in coding the possible form(s) of an **(unmodified) pivot** (or **subject** in **negative existential constructions**, broadening the typological patterns of articles within a particular language as well as crosslinguistically.

In many languages, the pivot (i.e. subject) in a negative existential construction or context differs in form from the pivot in an affirmative existential construction, as we can see in negative existential constructions in English for mass nouns.

(Affirmative) existential	Negative existential
(i) there is water in the bucket	(ii) a. There is no water in the bucket
	(ii) b. There isn't any water in the bucket

This is an area of great interest to syntax and semantics. It connects to the (very substantial) syntax-semantic literature on negation and **NPIs**(**N**(egative) **P**(olarity) **I**(tems)). It raises questions about the scope of indefinites under negation, their form(s), and the syntactic distribution of various expressions [3]

---

[1] Dipartimento di Comunicazione ed Economia, Università degli Studi di Modena e Reggio Emilia
http://personale.unimore.it/rubrica/dettaglio/cguardiano
[2] Department of Linguistics, UCLA
koopman@ucla.edu and linguisticexplorer@ucla.edu
[3] In the future, we intend to develop questions about the form, distribution and interpretation/scopal properties of indefinites objects in negative contexts
Please contact us if you are interested in collaborating some sswl properties about scope.

5.2 Expanding the current coding schema.

We build on our existing definitions and classification schema of the 3 properties *_1 can be bare, _2 must have an "article", _3 can have an article)*. These are not sufficient for capturing the known linguistic variation in negative existential contexts, as we detail below. We will:

1. broaden the notion of "article" in _2 and _ 3 to "article/specialized form" in negative contexts;

2. add an additional property _4 "an article disappears".

3. illustrate the settings for different languages.

4. And finally, we discuss how to deal with cases in which there are different strategies within the same language that cannot all be captured by the properties. In this case, our guidelines will be to code the strategy for the most neutral/ least marked negative existential strategy (see section 5.5).

5.3 There are two "articles" in French affirmative existentials, but only one in negative existentials.

We will start with (unmodified) indefinite mass nouns.
In French affirmative existential constructions, the pivot of a mass noun has two "articles" (*de* ART1, related to Genitive Case, and the definite article (*le/la/les* ART2), which combine as *du, de l', de la, des*, sometimes called "Partitive article").

(6) il y   a   de l'eau   dans le lac
     it there has DE LE.water in   the lake
     "There is water in the lake"

In affirmative existential constructions in French, bare (mass) noun phrases are not allowed. The value for _1 is NO. Since both "articles" are obligatory, the value for _2 is YES, and for _3 it is NO:

(Values for S02:
Indefinite mass noun in affirmative existential constructions for French:

_1 can be bare	_2 must have an art	_3 can have an art	found?
no	yes	no	French

In negative existential constructions, however, the two articles (*de* ART1 and the definite article (ART2behave differently: *de* (related to "Genitive") is obligatory, but the definite article *le/la/les* must be absent, i.e. ART2 cannot be present under negation.

(7) il (n')y   a   pas d'eau   dans le lac
     it (NEG).there have NEG DE.water in   the lake
     there is no water in the lake

(8) *il (n')y   a   pas   de l'eau dans le lac
     it N'.there have NEG DELE.water in the   lake
     there is no water in the lake

To capture this important property, the properties for negative indefinites must be further refined and expanded.

To this effect, we add a fourth property to the list, which captures the fact an article cannot be present or "disappears" on the pivot in negative existential constructions, and do so for each of the indefinite <mass, sg or pl count N > pivots.

> (9)   Property Subj_Neg.ex._4 an indef <mass/sg count/plural count> subject in a negative existential context: an article disappears

5.4   Indefinite Mass Nouns in negative existential contexts

This yields the following 4 properties for indefinite mass nouns in negative existential contexts, and their values for French:

_1... can be bare : NO
_2. .. must have an article: YES (*because "de" is obligatory*)
_3 .. can have an article: NO *because "de" is obligatory, (even though le cannot be present)*
_4... an article "disappears": YES (*in comparison to the affirmative existential, the definite article le disappears (i.e. must be absent)).*[4]

Indefinite mass nouns in negative existentials (*preliminary coding schema*):

_1 can be bare	_2 must have an art	_3 can have an art	_4 an article disappears	found?
no	yes	no	yes	French

_1 is no, and _2 remains yes, since *de* must be present in this context.

Because *de* is obligatory, _3 is no (since neither *de* not *le* is optional _3). Since _4 lacks an article is 'yes", it must be the case that there are two articles in the affirmative.

Other potential candidates for language like French, where an article goes missing in negative contexts are Bantu languages with so-called "augments". Note that currently, we don't make a distinction as to whether the article disappears obligatorily or optionally.

5.4.1   Broadening the notion of "article" in negative contexts: English NPIs and the negative determiner *no*:

Under _2 and _3, we include not just *de* (or genitive case which in some languages count as an "article"[5], but also negative determiners like English *no*, "special" articles like *any* in English, or other special indefinite forms, (reduplicated or not), often called Negative Polarity items, NPI).

It will be important to give as many glossed examples in the database as necessary.

---

[4] A language with no article in the affirmative, will automatically be NO for _4

[5] when it alternates with a different form like accusative for example

14

(10)   a.   There is no water in the lake
            b.   There isn't any water in the lake
            c. ??There isn't water in the lake

There seems to be some speaker variation on the acceptability of (10) c. The set value could go either way depending on the native speaker linguist. As a first step, this should be indicated in a comment that accompanies the example. In the future it may lead to having different varieties of "English".

(11)   Negative existential poperties:

         Property S05_2 and S05_3:
         A (mass/sg count/plural count) noun in a negative existential context must/can have an **article/ specialized** form: Yes/No

Table for Indefinite Mass Nouns in negative existential constructions(*final schema*[6]

_1 can be bare	_2 must have a art/specialized form	_3 can have an art/specialized form	_4 an art disappears	found?
no	yes	no	yes	French
no?	yes	no	no	English

5.4.2   French and Italian

French and Italian differ in that in Italian a mass nouns in the negative existential construction can either be bare, or have the partitive article (di + definite article)[7]. i.e. **the two articles alternate with the bare form, one article cannot disappear**. This means that Italian will have a YES value for _ 1, for _3, and a NO value for _4.

> This means that we determine the value for _4 w.r.t the forms in _2 and _3. Since the two articles must be both present _4 will be NO in Italian. In French, since the definite article is excluded in this context, _4 will be,YES.

Table for Indefinite Mass Nouns in negative existential constructions(*continued*)

_1 can be bare	_2 must have a art/specialized form	_3 can have an art/specialized form	_4 an art disappears	found?
no	yes	no	yes	French
yes	no	yes	no	Italian
no?	yes	no	no	English

5.4.3   Expanding the typology: more languages

Here we incorporate examples from a a few more languages of the form indefinite mass nouns as subjects/pivots in negative existential contexts, and show how they should be classified. The same needs to be done independently for singular and plural count Ns. We briefly return to singular count nouns in section 5.5.

---

[6] NB:we must code lacks an article w.r.t. the affirmative, as we can see from Croatian below

[7]In a affirmative existential construction in Italian, the mass noun must be bare, and the partitive article is excluded (differing from French in this respect).

5.4.4   Croatian: "Genitive of negation" (Slavic)

We start with the affirmative existential construction in Croatian. At first blush, Croatian does not really seem to have a special affirmative existential construction[8].

(12)   frizideru je voda
         in.fridge is water.nom
         there is water in the fridge
         # not: the water is in the fridge

(13)   Voda   je u frizideru
         water.NOM is in fridge
         the water is in the fridge *not*: there is water in the fridge
         comment: in this order the (mass noun) subject is interpreted as definite

However, since word order encodes the difference in interpretation, we treat **??** as an instance of a special existential construction, and test all properties of existential subjects in this order/context. **Note that the pivot carries nominative case**.

Negative existentials and the "genitive of negation".   Croatian does have a special form for negative existential predicates, consisting of a contraction of *negation 'ne'* and *'imati' 'have'*:

(14)   Nema   vode   u frizideru
         not.have water.GEN in fridge
         there is no water in the fridge

In negative existential contexts, the indefinite pivot must be genitive (and cannot be nominative). Since genitive contrasts with nominative in the affirmative construction, the genitive case is treated as an "article" (just like Romance *de/di*), according to our definitions for "articles".

Table for Indefinite Mass Nouns in negative existential constructions(*continued*)

_1 can be bare	_2 must have a art/specialized form	_3 can have an art/specialized form	_4 an article disappears	found?
no	yes	no	yes	French
yes	no	yes	no	Italian
no?/yes?	yes	no	no	English
no	yes	no	no	Croatian

Croatian differs from French for the values of _4. Since Croatian does not have an article in the affirmative, the value for property _4 "an article disappears" is *no*, because the language simply doesn't have an "article" that can disappear. (nominative will not qualify as an article, because the nominative form in the affirmative does not distinguish between definite and indefinite interpretations.) It is important to make sure that the value for _4 is based on a comparison with the *presence* of an article in the affirmative form.

---

[8]Data kindly provided by Daniela Culinovič

5.4.5 Mandarin Chinese

(15) Mandarin

        hu li   mei-you   shui
        lake inside NEG.PERF-have water
        There was no water in the lake.'

Table for Indefinite Mass Nouns in negative existential constructions(*continued*)

_1 can be bare	_2 must have a art/specialized form	_3 can have an art/specialized form	_4 an art disappears	found?
no	yes	no	yes	French
yes	no	yes	no	Italian
no	yes	no	no	English
no	yes	no	no	Croatian
yes	no	no	no	Mandarin

5.4.6 Gungbe

(Data provided by Enoch Aboh.)

Il y a pas d'eau dans le lac

Sìn   má   tò   tɔ̀   lɔ́   mè

Eau   NEG   être-LOC   étendue.d'eau   DET   P

Il y a pas pas d'enfants dans la maison.

Vǐ   má   tò   xwé   gbé

Enfant   NEG   être-LOC   maison   P

Il y a pas de medecin au village

Dòtó   má   tò   tò   lɔ́   mè

Médecin   NEG   être-LOC   village   DET   P

Property _4 is NO for Gungbe, because there is no article in the affirmative, i.e. there is simply no article that can disappear. Thus, Mandarin and Gungbe represent the same values for negative mass nouns in existential constructions.

Table for Indefinite Mass Nouns in negative existential constructions

_1 can be bare	_2 must have a art/specialized form	_3 can have an art/specialized form	_4 an art disappears	found?
no	yes	no	yes	French
yes	no	yes	no	Italian
no	yes	no	?*no*	English
no	yes	no	no	Croatian
yes	no	no	no	Mandarin
yes	no	no	no	Gungbe

5.4.7 Samoan (Polynesian)

Samoan[9] has *se*, a specialized NPI like determiner in negative existential constructions (and other contexts). [10]

(16) e   leai   se araisa i   totonu o le pou
      GEN NEG.exist SE rice   LOC inside of the bowl
      There is no rice in the bowl

In positive existential a mass noun pivot is marked with *le* .

(17) o lo'o   iai   le masima i   totonu o le   supu
      O PRES EXIST LE salt   LOC inside of the soup
      There is salt in the soup

Table for Indefinite Mass Nouns in negative existential constructions

_1 can be bare	_2 must have a art/specialized form	_3 can have an art/specialized form	_4 an art disappears	found?
no	yes	no	yes	French
yes	no	yes	no	Italian
no?	yes	no	no	English
no	yes	no	no	Croatian
yes	no	no	no	Mandarin
no	yes	no	yes (?)	Samoan
etc				

We see that Samoan and French come out in the same way, though by the definite article *le,la,les* disappearing in French, but in Samoan by the general article *l* alternating with *s l.e/ s.e* for mass nouns and singular count nouns. For plural count nouns, the D is zero, and the NPI Det is *ni*. Examples and comments should allow deducing this distributional difference. How to analyze this difference is ultimately a question for linguists, not a question of generating the cross linguistic patterns, which is the general purpose of these properties.

---

[9]Data from Mosel and Hovdhaugen, 1992 *Samoan Reference Grammar* and a 2 quarter UCLA field methods class on Samoan (2006/2007).

[10]The existential predicate appears to be composed of a locative "case" *i* and *ai* ('there'). the negative existential is composed of the negation *le* and *ai* "there".

5.4.8 Impossible languages

The way our properties are formulated define intrinsically impossible languages. For example (at least for each strategy within a language, see the discussion in 5.5):

(18)   a.   _1 "can be bare" and _2 "must have an article" cannot have the same values

   b.   _2 "must have an article", and _3 "can have an article" cannot have the same values

This is indicated in red under the double line in the table below.

Table for Indefinite Mass Nouns in negative existential constructions

_1 can be bare	_2 must have a art/specialized form	_3 can have an art/specialized form	_4 an art disappears	found?
no	yes	no	yes	French
yes	no	yes	no	Italian
no?	yes	no	no	English
no	yes	no	no	Croatian
yes	no	no	no	Mandarin
no	yes	no	yes (?)	Samoan
yes	yes			0
no	no			0
	yes	yes		0
	no	no		0

5.5   Singular Count nouns: negative existential contexts: indefinite articles, the numeral "one", and NPIs

French (unmodified) indefinite pivots in affirmative existentials require the indefinite article "un" (which is homophonous with the numeral *one*). In "neutral" negative existential constructions *de* appears (19) b): *de* is incompatible with *un* , as shown in (19) c.

(19)   a.   Il y a un médecin au village
            There is a doctor in the village

   b.   Il n y a pas de médecin au village
            There is not DE doctor in the village

   c.   *Il n y a pas d'un médecin au village
            There is not DE ONE doctor in the village

On the basis of these examples, we should set the values for singular indefinite count nouns in negative existential constructions as follows: _1 is NO (since the count noun cannot be bare), _ 2 is yes (since an article i.e. *de* is obligatory), and _4 as YES, since *un* disappears.)

Indefinite singular count nouns in negative existentials:

	_1 can be bare	_2 must have a art/specialized form	_3 can have an art/specialized form	_4 an art disappears	found?
de	no	yes	no	yes *'un' disappears*	French

This takes care of "neutral" (or least marked) negative existential constructions: but there are other strategies as well

It is sufficient for our purposes for the setting of this property to consider the most natural or "neutral" negative existential constructions in case of conflicts.

5.5.1 More than one strategy.: When 'un' does not disappear, and more about the possible appearance of 'aucun"

The problem is that languages may have more than one strategy in negative existentials. In French, the indefinite article *un* cán appear as well, particularly when it is stressed i.e. *FR "UN"*. Note that in these cases, *de* cannot appear, and a special interpretation (*pas un seul* 'not a single' arises:

(20) a. Il n y a pas UN médecin au village
    There is not ONE doctor in the village
  b. There isn't (even) a (single) doctor in the village

The same applies to English *one/a* , when it interpreted as "(not) a single". Stressed ÚN and *ONE* are indistinguishable from numerals, and seem to be interpreted close to *not (a single) one*. This raises complex questions about the relation of the unstressed indefinite article *un* and the numeral *one*: are these different categories or not?
For our purposes here, we put the question of how to resolve this problem here aside. In case of conflict it is sufficient to consider what happens in neutral negative existentials. Adding examples and comments to the database when such cases arise, will be useful for further development
We clarify the problem further by examining how we would we set _4 for indefinite count nouns on the basis of the examples and interpretation with (20-a) under negation. We add a special line for the properties of French w.r.t. to *un*, and note the problem w.r.t the non-neutral status here.
(ideally we should code the inventory of the different forms in French in the database, as we do in the table below.

	_1 can be bare	_2 must have a art/specialized form	_3 can have an art/specialized form	_4 an art disappears	found?
de	no	yes	no	yes *un*	French
un	no	yes	no	?no/?yes "special reading?'	French

(21) Il n' y a aucun médecin au village
   There is AUC.UN doctor in the the village
   NB:*pas* must be absent

On the basis of this example, we can add another line for *aucun*, which clearly is build on *un, une*, hence is NO for _4 . Thus, different strategies may (but do not have to) lead to different codings. This is why we restrict the coding for the moment to the

least marked strategy. Examples for the other strategies can always be added, and commented on.

	_1 can be bare	_2 must have a art/specialized form	_3 can have an art/specialized form	_4 an art disappears	found?
de	no	yes	no	yes *un*	FR
un	no	yes	no	?no/?yes "special reading?'	FR
aucun NP	no	yes (aucun)	no	no "un is part of aucun'	FR

> Important: with these indefinite counts nouns *de, un, or aucun* in negative existentials the (unmodified) indefinite count noun pivot must be interpreted under the scope of negation (i.e. the meaning can be paraphrased with *it is not the case that there is a doctor in the village, i.e. there are no doctors in the village*). These sentences cannot be interpreted with *'one' or 'un'* scoping above the negation. This can be concluded from the fact that these sentences cannot mean: ** there is a single doctor who is not in the villlage* which would be true if all other village doctors were actually in the village.

We return to the (possible) form(s) and interpretation of indefinite count noun objects under negation in the near future.

This document discusses how to develop a coding schema, and how to formulate the queries so as to capture the variation found in natural languages. The next step is to gather the data from individual languages to code this variability in a systematic fashion: it will be important to give examples. The search function (cross) will allow extracting all patterns in the database for these 4 properties, and save and download the results.