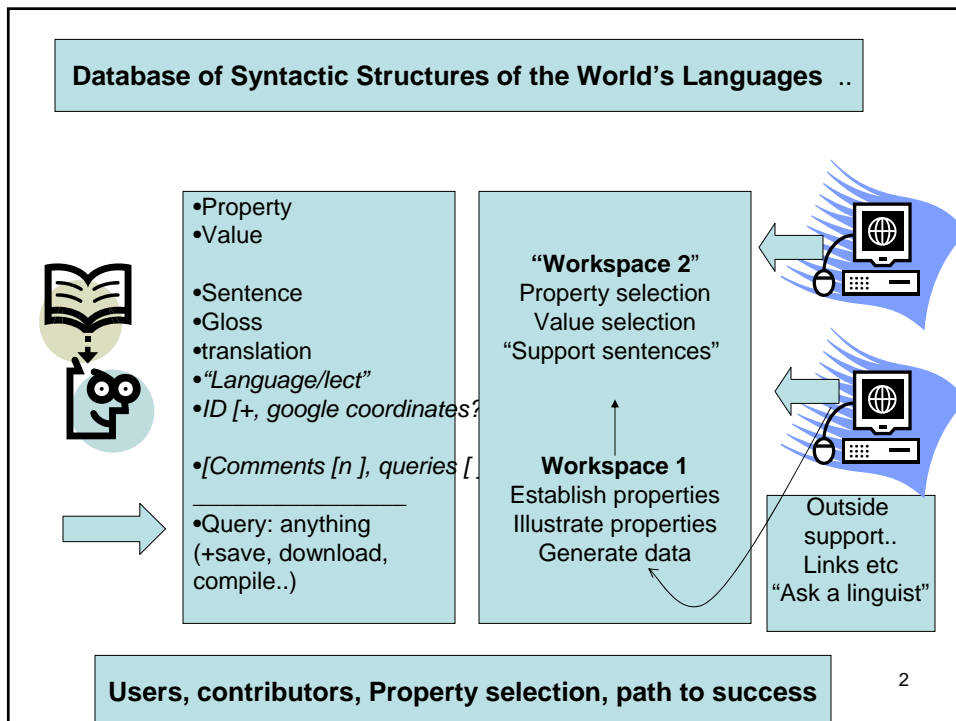


SSWEL nov 10 2007
Hilda Koopman
UCLA

*Some comments following up on the first
and second round of comments..*

1



The need for a tool of this kind..

some additional remarks (in addition to Richie's talk)..
Necessary for future progress.

+where is current progress in the field coming from?

In part from “Microparametric”
work/cartography.

In addition to the sources sources
mentioned at workshop (northern Italian
dialects)

- Meertens Institute: SAND, MAND,
(mapping property by zipcode), **Edisyn**
(european dialect syntax) (see next slide
for links)

3

Links

- Sand (Syntactic Atlas of Dutch dialects)
<http://www.meertens.knaw.nl/projecten/sand/sandeng.html>
- Mand (morphological atlas Dutch dialects)
<http://www.meertens.knaw.nl/projecten/mand/>
- Edisyn (European Dialect Syntax):
<http://www.meertens.knaw.nl/projecten/edisyn/>

4

The tension between SSWEL for "well studied
"languages" and "unknown languages"
the ultimate goal are data on individual grammars

What we (can) ask from SSWEL is not the same, depending on the starting point):

1. example of what we can ask of "languages" that have been well described;
2. of what we can ask or can hope to ask of "languages" that have not yet been well described

and the problem of sociolinguistic factors (observable in the history of the field): from properties of "Dutch" to finer grained properties of individual grammars is a bit from going to "we speak the same language therefore we have the same judgments patterns/internal grammars", to acceptance of individual variability as a true fact and an important source of syntactic theory. (this took a lot of time!)

and the problem of syntactic atoms.. (much tinier than usually assumed)

5

An example of the type of questions we can ask of languages for which we have identified areas of individual variation within the general "language" area:
...verbal clusters/complex V formation/restructuring in West Germanic)

..inspired by Koopman and Szabolcsi (2000), in which work

- ..we makes specific predictions about the grammars of individual speakers in a quite large domain (which includes auxiliaries, modals, infinitives, participles, particles, directional PP, incorporated nouns, nominal small clauses and adjectives). (i.e. predictions about patterns of judgments (positive and negative) of individual speakers)
- We need to test whether these predictions hold or not! (modeling knowledge of individual speakers) but we cannot, because we don't have the data...
- **what are these patterns?** → requires massive data collection of the behavior of individual speakers.
- Our project-- > do "controlled internet fieldwork" to gather lots of data ("judgments questionnaires") (ongoing pilot study work Cornips and Koopman (Meertens Institute).)
- Would be important to incorporate this type of project with SSWEL
- General message: we need to find out about individual grammars

6

describing “little known” language X

- the initial effort will have to be more toward documenting lots of properties of language X (while keeping track of identity of individual speakers). Use what we have learned theoretically, typologically etc to gather these data (i.e. this requires a rich data gathering environment, geared towards hierarchies and cartography, guided by questionnaires which generate “prompts”)
- Involve native speakers linguists (requires structured workspace to generate examples)

7

- Use “drop down” menus (for data gathering on little known languages); (prompts, etc”) questionnaires
- Use “drop down” menus for exploring further syntactic questions; import standard tools of the trade: (“labtests”)
- *for example: a speaker is interested in trying to find out properties of a substring or a discontinuous string x and y” in language x. What kind of tests are there to ask questions about this??*

Drop down menu: coordination, questioning, relatives, focusing, passivization, raising, fronting, scrambling, ellipsis, general c-command tests.. association with “only”, “also”, or other particles,.., pronominal binding, different types of ellipsis, scopal interactions, what can intervene, ordering, reconstruction, etc.....)

This will yield a lot of data (though not immediately analyzed in finely grained property,value pairs, but searchable by id), and might lead to speakers “forming” questionnaires in specific areas of inquiry.

8

towards ensuring data quality

- Ensuring the quality of the data...
- Properties/values should be verifiable...
(examples illustrating a “property” “value”).
--support sentences--
and justified (for new property/value)
- requires a rich “work place”, “a lab” which provides the tools to determine the properties of a particular string.
- and a place where properties are justified and illustrated (workspace..)

9

The Path to success SSWEL: “addiction”

- How to create an “addiction”?
→ ensure the quality of the data;
- Make it into an indispensable research tool which is fun (=addictive) to use and fun(=painless) for datacreation.
- Once the “ingredients for data entering environments ” are set up, the data gathering/entering part should be (relatively) fun, and worth the effort of what the user can get out of it..
- Use the lightest structure possible (avoid moderators), let users generate “questions” in the database, maybe create “chatrooms”, working groups..
- In every aspect: Ensure superlative user friendliness (“minimize frustration”)

10

A start of a toy example of a user (hk): and what it requires from the database (properties have to be finely grained and broken down).

Underlying question: in some languages imperative morphology is morphological identical to participles or past tensed forms. What other properties does this correlate with?

- Imperatives:

- Query:

- Search for all *Languages/IDs* for which the form of the IMP(erative)=
“**past participle**”
- & addressee is covert; (*lit. “carried me”*)
- (& {order V.imp.1sO &..1sO.V-T(root decl)}) (*porte-moi, il me porte*) etc)
- (& , OR, NOT) , & inchoative= value x, &causative of vPs= y))

- Problem: what is meant by past Participle (“glossing and the naming problem”);

Depending on the language, the following reasonable glosses will come up:

- { PAST.PART, PAST, PERF, COMPL, FACTATIVE,..} and {not glossed}.
- The query should be formulatable in such as way as to encompass these glosses, (except for ‘not glossed’).

11