Universität Zürich, 2.12.2021
Framework-free grammatical theorizing (1)

Martin Haspelmath
Max-Planck-Institut für evolutionäre Anthropologie (Leipzig)

1. Overview

– frameworks as (partly) **isolated communities**

– *theories* vs. *methods* vs. *data*

– a GENERAL FRAMEWORK must be either a method or a theory (it cannot be both simultaneously)

– general linguistics and particular linguistics (g-linguistics/p-linguistics):
> particular analyses can contribute directly to g-linguistics only if the framework is a theory, i.e. if it is **hypothesized to be innate**

– the "diagnostic-fishing" problem:
> linguists often use **different** (cherry-picked) **diagnostics for different languages**

– the proposed solution:
> – (general linguistics:) use **the same criteria** for all languages
> – (particular linguistics:) use categories **appropriate for the language**

2. Frameworks as isolated communities

Syntacticians (and morphologists) tend to inhabit isolated communities.

An excerpt from a book manuscript review (for Oxford University Press) of Stefan Müller's book *Grammatical theory* (Language Science Press, 2016).

My final point is that it's difficult for me to see what audience the book would attract. Most researchers in syntax are wedded to one particular theory and the type of overview offered in this book would not be of great interest. Of course, some of them, may want to take the occasional view across the fence, but my feeling would be that they might rather pick up a overview article on Minimalism or TAG written by an insider instead of the present book.

https://twitter.com/LingMuelller/status/1462036336798535685

The isolated communities are "gated" ("view across the fence"), and they are created by teaching courses on particular framworks.

Adger (2021):

> "I accept that not everyone who is analysing the syntax of languages *using* a particular theory might be motivated by the broad aims I set out above. They may just be interested in a particular phenomenon in their language, and they might find that **the theory they learned in graduate school** is useful for that task."

Corey Wright
@coreywrite

"people working in the dominant paradigm literally don't have time to learn HPSG or LFG, plus those approaches are wrong anyways" is a real take i just saw on the bird app!

"We are all constructionists": https://dlc.hypotheses.org/2470
"We are all structuralists": https://dlc.hypotheses.org/2356
"How formal linguistics appeared and disappeared": https://dlc.hypotheses.org/1698

3. Theories – methods – data

Linguists often use peculiar terms that are not current in other sciences:

> – "argue that..."
> – "committed to ..." (e.g. Lakoff's 1991 "cognitive commitment")
> – "account for ..."
> – "analyze X as Y"   (e.g. "analyze the English word *to* as an auxiliary")

Linguists often use "theoretical" when they mean "general", and "descriptive" when they mean "not very prestigious". And linguists never know what they mean by "explanation" or "explanatory".

This seems to be a problem.

So let us take a step back.

Science consists in the collection of **data** in order to test **hypotheses** (as well as the creation of **theories** from which new hypotheses can be derived).

Scientists apply a range of specific **methods** to their data in order to test these hypotheses.

Kabatek et al. (2021: 3):

> new complexity to this picture, the present volume brings DOM back to its origins, the Romance languages, and aims to deliver fresh insights arising from new data, new methods, and new theoretical approaches.

A **theory** is a coherent body of hypotheses that are subject to empirical testing. The more a hypothesis resists disconformation, the more it is likely to be true. A rich theory allows one to derive a rich set of testable hypotheses.

As a mass noun, "theory" seems to mean roughly the same as "theorizing", e.g.

> "Linguistic theory is concerned primarily with an ideal speaker-listener..."

So Haspelmath's (2010) "framework-free grammatical theory" really means "theorizing about grammatical phenomena without a general framework".

> "Theoretical linguistics" contrasts with *applied linguistics*, so "theorizing" really refers to everything that is not data gathering (cf. Haspelmath 2021a).

I grew up as a linguist without a framework-based community
>   (1983-85: Vienna, Wolfgang Dressler: "Natural Morphology" as a framework??)
>   (1985-86: Cologne, Hansjakob Seiler: "UNITYP" as a framework??)
>   (1987-88: Buffalo, Joan Bybee: "Grammaticization" as a framework??)

So I always asked myself what creates communities and keeps them alive.

> [See also "The "typology vs. theory" mistake": https://dlc.hypotheses.org/1915; "typology" should not be understood as a community.]

4. On (general) frameworks as theories

– A framework is a coherent body of **concepts** that can be used to formulate hypotheses and theories.

In linguistics: *Frameworks* are often called "theories", e.g.

> Alexiadou: "The Principles and Parameters theory..."
>   "The Minimalist Program (MP) ... is a strong lexicalist theory."

> Legendre: "Optimality Theory..."
>   "structural descriptions expressed in the format of any substantive theory of syntactic representation (e.g. brackets, syntactic categories, etc.)

– Calling a framework a "theory" is justified if it is hypothesized that
the framework is innate; because scientists have long made hypotheses about what is innate and what is learned.

> Baker & McCloskey (2007: 286)
> "[absolute universals] ... must either be built into the design of the theory, or the theory must be developed in such a way that it guarantees their truth."

I have called this the **restrictivist approach** to comparative grammar (Haspelmath 2014). The idea is that the framework is **restrictive** in that it can only describe those languages that are actually attested.

In this view, ***"describing" and "explaining" is the same*** if the the description makes use of the elements of the framework/theory.

E.g

Universal generalization: OV languages lack question-word fronting

Kayne (1994: 54):   "antisymmetry framework":
• Heads always precede complements underlyingly:

V-O, COMP-clause

• Other orders must be derived, by leftward movement:

(1)   [*Yooko-wa Masa-o aisite iru*]S [[*to*]COMP [*t*]S]
[*Yooko-wa   Masa-o   aisite   iru   **to***]
Yoko-TOP   Masa-ACC   loving is   COMP
'that Yoko loves Masa'

[*Yooko-wa Masa-o aisite iru*]S   [[*to*]COMP   [*t*]S]
<   [[**to**]COMP   [*Yooko-wa Masa-o aisite iru*]S]

• The pre-complementizer position
('specifier of complementizer')
is assumed to be the position to which
question-words are fronted

• Since the Spec-Comp position is filled by the clause,
it cannot be the target of question-word fronting.

Hence: OV languages with final complementizers cannot
front their question-words (Kayne 1994).

Particular analyses can contribute directly to g-linguistics only if the framework
is a theory, i.e. if it is **hypothesized to be innate** (Haspelmath 2021a).

If each framework-based description/analysis is regarded as a **test** of the framework,
then the approach is coherent – the hypothesis is that the basic building blocks are
innate, and each new language (or each new phenomenon) is a test of the theory.

This explains the **"re-describing" nature** of many papers:
(i) one part describes the phenomena in a framework-neutral way
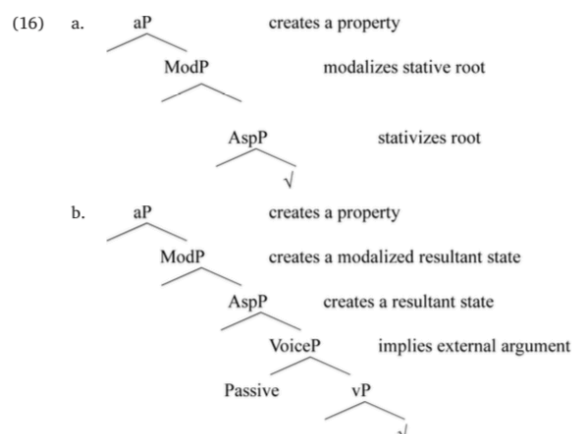(ii) a later part describes the same phenomena in a framework-bound way

e.g. Alexiadou (2018), on abilitive adjectives and experiential verbs in English and
Greek:
1. Introduction
2. Background (classes of psych verbs, …)
3. -Able adjectives and psych verbs
3.1. English
3.2. Greek
4. Towards an analysis
5. Conclusion

(32) a. John is easily satisfiable. John is easy to satisfy. John satisfies easily.
     b. John is so annoyable. John is easy to annoy. John annoys easily.

(40) To vivlio ine metafrasimo   apo ebirus   metafrastes.
     the book is translatable.NEUT by experienced translators
     'The book is translatable by experienced translators.'

(41) To vivlio ine metafrasimo   mesa se deka meres.
     the book is translatable.NEUT with in 10   days
     'The book is translatable in 10 days.'

(16) a. aP   creates a property
             ModP   modalizes stative root
                 AspP   stativizes root
                     √

     b. aP   creates a property
             ModP   creates a modalized resultant state
                 AspP   creates a resultant state
                     VoiceP   implies external argument
                 Passive   vP
                              √

So there are two modes of description:
     (i) generally accessible;
     (ii) framework-bound – and only the latter is called **"analysis"**.

The former is sometimes called "data", although it does not present the raw data –
syntacticians do not normally publish their **raw data** at all. It is also sometimes said to
present the "descriptive generalizations", contrasting "description" with "analysis."

When a linguist says that they "analyze phenomenon X as Y", they often mean
that X is a generalization formulated in language-particular terms,
and Y is an innate building block.

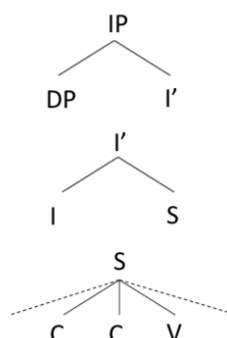Börjars (2021; Zurich lecture notes), on universal building blocks in LFG:

c-structure

Universal set of phrase structure rules from which a language chooses a subset. Some examples:

IP → DP I'

I' → I S

S → V C*

f-structure: attributes (features) and values

A feature value matrix: an unordered set of feature-value pairs

Types of attributes and their values:

Functional features	NUM PERS TENSE ...	Atomic values
Semantic feature	PRED	Semantic form
Grammatical relations	SUBJ OBJ ...	f-structure

F-structure is reasonably invariant across languages

(3) a. *I **played** football yesterday.*
  b. *Ik **heb** gisteren voetbal **gespeeld**.*

different c-structure, same f-structure

Legendre & Lutken (2021, Zurich lecture notes):

The difference between the two languages is understood as the resolution of a conflict among **two universal constraints**:

- SUBJECT (~ EPP; Chomsky, 1982): Clauses have subjects in SpecTP
- FULL INTERPRETATION (FULLINT): No expletives

based on Grimshaw and Samek-Lodovici (1998)

On the other hand, if the general framework is a **method**, it cannot contribute directly to g-linguistics.

5. On general frameworks as <u>methods</u> for description

A general framework can also be thought of as a *method for description*. For example, some linguists regard HPSG as an elaborate set of **notational conventions**, which are not restrictive but can express a wide range of generalizations.

Universal Dependencies (de Marneffe et al. 2021) is explicitly designed to be a notational framework, facilitating comparison across languages.

Many generative linguists seem to think of general frameworks as methods, as seen in the Adger (2021) quotation:

> "I accept that not everyone who is analysing the syntax of languages *using* a particular theory might be motivated by the broad aims I set out above. They may just be interested in a particular phenomenon in their language, and they might find that **the theory they learned in graduate school** is useful for that task."

But this is weird, because generative theories such as Minimalism or Optimality Theory are not good methods for description (they may be good theories).

On (general) frameworks as scaffolds (Haspelmath 2020):

A scaffold is a temporary structure that aids in constructing something bigger that has its own (and possibly unique) frame. Once the construction is finished, the scaffold can be taken away.

Typological concepts are often regarded as highly useful in helping language describers to find the correct generalizations in their language, because we find simlar factors operative in different languages.

Kabatek et al. (2021):

> What, then, is the third wave like? In Romance linguistics, there are two main directions for recent studies on DOM. On the one hand, researchers go back to apparently established issues and take a closer, more fine-grained look at the data. Typological generalizations have opened up new perspectives, but in doing so they have also broadened the horizon towards related phenomena. The interplay of different means of object differentiation within the same language is studied (e.g. clitic doubling or "indexing" vs. traditional "flagging" DOM, i.e. marking by an element that immediately stands by the object). Not only objects are considered; instead, whole constructions are also taken into account. And newly identified factors such as agentivity, telicity and affectedness are considered when looking closely at language variation. On the other hand, new methodological perspectives have been opened. In historical studies, the availability of large corpora makes it possible to work with comprehensive databases and

6. On particular frameworks,
i.e. language-particular descriptive notions

Each language has its own categories, and each language should be described in its own term.

This is an essential insight from a century ago (Boas, de Saussure, Sapir, Jespersen, and others). Languages should not be put into a Procrustean bed of concepts from Classical grammar ("traditional grammar").

e.g. Nick Evans:

> "[Franz] Boas made explicit many of the tenets that have become axiomatic in the best descriptive work: the importance of **describing each language and each culture on their own terms** rather than importing inappropriate European models, the need to discover **the inner design of each language** inductively through the study of texts, and the scientific responsibility to produce as undistorted a record as possible by setting grammar alongside a comprehensive dictionary and text collection" (Evans 2010: 36)

The **uniqueness** of the language-particular generalizations and structural oppositions is often associated with "structuralism", and indeed, linguists began to emphasize synchronic grammatical structures in the 1920s.

How to compare the particular structures of different languages is not immediately obvious, and many structural linguists were skeptical about large-scale comparison. Many thought that we first need to understand the bias induced by our habits of using European-derived concepts
                (e.g. Whorf, who coined the term "Standard Average European").
For example, the concepts of "subject" does not work well for ergative languages, because the "intransitive subject" is coded like the "transitive object".

Lezgian (a Dagestanian language):

(2)   Алиди киц кьена
        Ali-di   kic'   q'e-na.
        Ali-ERG dog   kill-AOR
        'Ali killed the dog.' (Haspelmath 2019)

7. The diagnostic-fishing problem

Many linguists assume that concepts such as "subject" or "clitic" or "serial verb construction" are concepts of general linguistics, but can be identified by **different diagnostics in different languages**.

But which diagnostics do we choose?

Diagnostic-fishing for "clitic":   Haspelmath (2015)
                "serial verb construction":   Haspelmath (2016)
                "subject":   Haspelmath (2019)
                "existential clause"   Haspelmath (2021d)
                "word"   Haspelmath (2011)

Table 1. Nine studies that examine wordhood using test batteries

	Zwicky & Pullum 1983	Kanerva 1987	Bresnan & Mchombo 1995	Ackema & LeSourd 1997	Monachesi 1999	Harris 2000	Milićević 2005	Lieber & Scalise 2006	Bickel et al. 2007
Free occurrence				+			+		
External mobility and internal fixedness	+			+	+	+			
Uninterruptibility				+					+
Non-selectivity	+	+			+	+	+		+
Non-coordinatability			+	+	+		+	+	+
Anaphoric islandhood				+				+	
Nonextractability				+				+	
Morphophonological idiosyncrasies	+	+			+	+	+		
Deviations from biuniqueness									+

Diagnostic-fishing is not a rigorous method for comparing languages, and it does not contribute to enlightening language-particular analyses.

It contributes to general linguistics only if the universal categories are **hypothesized to be innate** and if the hypothesis is confirmed again and again.

8. Conclusion

• the proposed solution:
  – (general linguistics:) use **the same criteria** for all languages
  – (particular linguistics:) use categories **appropriate for the language**

• typological comparative concepts need to be **rigorously defined** in a language-independent way (e.g. "affix" in Haspelmath 2021b), and ideally in a standardized way (Haspelmath 2021c)

• particular linguistics should describe each language **in its own terms**, not by means of typological concepts, though it can be inspired by them (Haspelmath 2020)

• a set of comparative concepts can be a general framework seen as a **method**

• if a general framework is intended as a theory, i.e. if its building blocks are hypothesized to be innate, then they should be tested systematically (rather than "taught in graduate school")

Finally: Some fundamental thoughts about science

– science begins when belief is undermined by doubt
– good scientists always ask themselves if they are perhaps wrong
– science is organized skepticism (Robert Merton)

References

Adger, David. 2021. On doing theoretical linguistics: A reply to Haspelmath. *Theoretical Linguistics* 47(1–2). 33–45. (doi:10.1515/tl-2021-2003)

Alexiadou, Artemis. 2018. -Able adjectives and the syntax of psych verbs. *Glossa: A journal of general linguistics*. Ubiquity Press 3(1).

Baker, Mark C. & McCloskey, Jim. 2007. On the relationship of typology to theoretical syntax. *Linguistic Typology* 11(1). 285–296. (doi:10.1515/LINGTY.2007.023)

de Marneffe, Marie-Catherine & Manning, Christopher D. & Nivre, Joakim & Zeman, Daniel. 2021. Universal Dependencies. *Computational Linguistics* 47(2). 255–308. (doi:10.1162/coli_a_00402)

Evans, Nicholas. 2010. *Dying words: Endangered languages and what they have to tell us*. Hoboken, NJ: Wiley.

Haspelmath, Martin. 2010. Framework-free grammatical theory. In Heine, Bernd & Narrog, Heiko (eds.), *The Oxford Handbook of Linguistic Analysis*, 341–365. Oxford: Oxford University Press.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80.

Haspelmath, Martin. 2014. Comparative syntax. In Carnie, Andrew & Sato, Yosuke & Siddiqi, Dan (eds.), *The Routledge handbook of syntax*, 490–508. London: Routledge. (https://zenodo.org/record/344909)

Haspelmath, Martin. 2015. Defining vs. diagnosing linguistic categories: A case study of clitic phenomena. In Błaszczak, Joanna & Klimek-Jankowska, Dorota & Migdalski, Krzysztof (eds.), *How categorical are categories? New approaches to the old questions of noun, verb, and adjective*, 273–304. Berlin: De Gruyter Mouton.

Haspelmath, Martin. 2016. The serial verb construction: Comparative concept and cross-linguistic generalizations. *Language and Linguistics* 17(3). 291–319. (doi:http://doi.org/10.1177/2397002215626895)

Haspelmath, Martin. 2019. Ergativity and depth of analysis. *Rhema* 2019(4). 108–130. (doi:10.31862/2500-2953-2019-4-108-130)

Haspelmath, Martin. 2020. The structural uniqueness of languages and the value of comparison for description. *Asian Languages and Linguistics* 1(2). 346–366.

Haspelmath, Martin. 2021a. General linguistics must be based on universals (or nonconventional aspects of language). *Theoretical Linguistics* (https://ling.auf.net/lingbuzz/005158)

Haspelmath, Martin. 2021b. Bound forms, welded forms, and affixes: Basic concepts for morphological comparison. *Voprosy Jazykoznanija* 2021(1). 7–28. (doi:10.31857/0373-658X.2021.1.7-28)

Haspelmath, Martin. 2021c. Towards standardization of morphosyntactic terminology for general linguistics. In Alfieri, Luca & Arcodia, Giorgio Francesco & Ramat, Paolo (eds.), *Linguistic categories, language description and linguistic typology*. Amsterdam: Benjamins. (https://ling.auf.net/lingbuzz/005489)

Haspelmath, Martin. 2021d. Existential constructions and two types of comparative concepts: Construction-functions and construction-strategies. Paris. (https://zenodo.org/record/5728643)

Kabatek, Johannes & Obrist, Philipp & Wall, Albert. 2021. The third wave of studies on DOM in Romance: An introduction to this volume. *Differential Object Marking in Romance*, 1–20. Berlin: De Gruyter. (doi:10.1515/9783110716207-001)

Lakoff, George. 1991. Cognitive versus generative linguistics: How commitments influence results. *Language and Communication* 11(1/2). 53–62.

Müller, Stefan. 2016. *Grammatical theory: From transformational grammar to constraint-based approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press. (http://langsci-press.org//catalog/book/25)