Caroline Heycock, Edinburgh, PI
Jutta Hartmann, Bielefeld, PI
Isabelle Roy, Nantes, Co-I
Roberto Zamparelli, Trento, Co-I

The syntax of nominal copular clauses: theoretical and empirical perspectives (abbreviated project proposal)

1 Context

1.1 Introduction

At least since the work of Aristotle, there is wide agreement that all sentences have a subject and a predicate; predication is a fundamental aspect of human language. Predicates can be divided into those that are built around a verb ("verbal predicates") and those which are not ("non-verbal predicates"), and it is a question of long-standing interest in linguistics and philosophy how these two kinds of predicate differ, and what the differences entail for the ways in which sentences are constructed and how their meanings are put together from the meanings of their component words. Of the two types of predication—verbal predication and non-verbal predication—the former is much better studied and understood than the latter. In this project, we aim to shed light on a central type of non-verbal predication—predicates built around a noun, rather than a verb ("nominal predicates")—through a systematic investigation that is innovative both in its breadth and its depth. The project will bring together both theoretical insights from detailed studies and a broadening of the empirical landscape beyond well-studied languages to less studied ones. This will yield a better understanding of the principles underpinning the syntactic and semantic characteristics of non-verbal predication, and also, importantly, provide tools for other researchers to further extend work on these fundamental questions about the nature of human language.

Nominal predication involves combining a referential or quantified expression (the subject) with a nominal predicate, as in (1). In a number of languages, including English, a copular verb is also required, but even in such languages there may be contexts where the copula can be absent—see the contrast in (1a) vs. (1b).

(1) a. The speaker is an expert.
b. I consider the speaker (to be) an expert.

Already it should be noted that although the canonical nominal predicate is indefinite, some definites, for example superlatives, have a predicative use:

(2) a. The speaker is the most famous expert.
b. I consider the speaker (to be) the most famous expert.

Copular clauses of the kind illustrated in (1) and (2) are referred to in the literature as predicative or predicational; we will use both terms in what follows.

While it is a longstanding assumption that the copula observed in (1) and (2) is essentially semantically vacuous, a major puzzle recognized early on in the philosophical literature is that what looks like the same copula can also occur with two referring expressions, rather than requiring the second to be predicative. This is characteristic of identity/equative/equational sentences such as (3) (we will henceforth use the term *equative*).

(3) You can't have met Stephen King but not Richard Bachman!
Stephen King **is** Richard Bachman!

The equative copular clause in (3) appears symmetrical, in the sense that the two noun phrases *Stephen King*/*Richard Bachman* could equally well appear in the reverse order, with no apparent change of meaning. This is exactly as expected given that in the language of logic, equation/identity is an entirely symmetric relation. Strikingly, however, it seems that the symmetry of

examples like (3) is the exception in human language, rather than the rule. For example, the expressions of "mistaken identity" in (4) are not symmetrical, as the two sentences do not have the same meaning:

(4) a. In the dark, I thought you were your mother.
b. In the dark, I thought your mother was you.

The mismatch between the symmetry of equation in logic on the one hand, and the frequent asymmetry of equative sentences in natural language on the other, is then a further puzzle.

Yet another puzzle arises in cases like (5). The postcopular phrase in A's reply in (5) appears to be a referring expression (deictic expressions and names resist predicative use). Given the types of copular clause mentioned so far, that would make A's reply an equative. But there is evidence that here the *first* nominal is not referring in the usual way—one indication is that it pronominalises with the neuter pronoun *it*, as in B's reply. *It* is normally infelicitous if the referent is a human:

(5) Who won the marathon?
A: I think the winner is that woman in blue/Anne Smith.
B: No! I'm pretty sure it's/#she's the woman in red/Beate Jones.

This third type of copular clause is what Higgins (1979) named "specificational."

Even though a distinction between equative and predicational copular clauses has long been recognised (a classic reference is Russell 1905), and specificational clauses have attracted significant interest since Higgins' work in the 70s, there is comparatively little systematic cross-linguistic work on the syntactic properties of nominal copular clauses, particularly those with two potentially referential DPs, even though, as will be outlined below, it has become clear that there are (morpho)syntactic distinctions that at least partially correlate with the different interpretations.

Relatedly, even though there has been substantial progress in the understanding of the internal structure of indefinite noun phrases (see e.g. Roy 2013), and the semantics of definites (see e.g. Schwarz 2009, Coppock and Beaver 2015), most of the attention on definite noun phrases has been on noun phrases in argument (non-predicate) positions, and there has been little systematic application to nominal copular clauses.

The aim of this project is to address these gaps in the study of predication, both theoretically and empirically, by investigating the syntactic and semantic properties of the subtypes of nominal copular clauses illustrated in (2)–(5), considering both the structure of the clausal environment and the internal structure of the noun phrases involved, and extending detailed and systematic crosslinguistic comparison to a wider range of languages. Based on an overview of the state of the art in 1.2 to 1.5, we will present our more specific research questions in 1.6.

1.2 Classification

As stated above, the study of nominal copular clauses has long recognized a difference between equative and predicational copular clauses, with Higgins subsequently adding a third, specificational type (for now we set aside Higgins' fourth "identificational" category, restricted to examples with deictic pronouns in subject position, such as *This is Susan*.)

Since Higgins' seminal work, there have been a number of attempts to reduce these three different types to two, or even one single underlying structure (see den Dikken 2017 for an overview), typically by taking the specificational sentences to be syntactically derived from predicationals by a process of inversion (see Heggie 1988, Mikkelsen 2005), or instead, to be a subtype of equative clauses (Heycock and Kroch 1999, Jacobson 1999). There have also been claims that all putative equatives are in fact predicational or specificational (see e.g. Moro 2006). Conversely, there have been arguments for more fine-grained distinctions. Declerck (1988), for example, suggests an additional, definitional type. More recently, Roy (2013) proposes distinguishing two different types of nominal predicative sentences (defining and characterizing) correlating with the functional heads present in the predicate nominal. Additionally, as mentioned above, there are a number of cases

discussed in Heycock (2012) where both DPs appear to be referring expressions, but which are clearly asymmetric in interpretation (i.e. exchanging the two DPs leads to a different meaning), such as hypothetical cases (*if I were you*), role playing (*Sean Connery is James Bond*) and dream contexts (*In her dream, she was Michelle Obama*); a further case that has attracted attention in the philosophical literature but has been only briefly discussed by linguists (Schlenker 2003) is that of mistaken identity as illustrated in (4) (*In the dark I thought you were your mother*). For convenience we can group all these cases together under the label "counterfactual equatives." To the best of our knowledge there is no work that investigates the syntax of these cases at any level of detail, or even whether or not they form a single class syntactically. Questions concerning (a)symmetry in the syntax of equatives parallel corresponding semantic and philosophical questions, some of which are being addressed in the one-year project "L'identité dans le langage et dans la pensée" by Co-I Roy and collaborators in Nantes, which started in January 2023. The results of this project will be integrated into our investigations.

1.3 Differences in grammatical properties

Predication as property ascription is a logical/semantic relation, and one fundamental question that will be addressed in this project is whether there is a unique syntactic counterpart to it. One aspect of this overarching question relates to the classifications just discussed. While the starting point for these classifications is usually the differences in interpretation, there are a number of properties that correlate at least in some languages with these semantic types. For English the differentiation between predicational clauses on the one hand and equative and specificational types on the other has been argued to correlate with the possibility to occur in **small clauses**:

(6) a. Mary considers [that {an island/*?Schiermonnikoog}] (Heycock, 2013, 341)
b. *I consider [the culprit John] (Moro, 1997)

However, there is often a confound arising from the fact that *consider* imposes the requirement that the embedded predication involve a subjective judgment, which might for example already explain the contrast in (6a). Further, careful investigation of different types of "small clause" is required, given that what looks like the "same" small clause that occurs in the ungrammatical (6b) is grammatical—and seems to get a specificational interpretation—after *make*, when used in an epistemic sense as in (7) (Heycock, 1994a).

(7) But if what Jim says is true, that makes [the culprit John]!

Thus we need to investigate the possibility of different kinds of hidden structure in small clauses.

A second phenomenon that shows up potentially syntactic distinctions between some of the types under discussion is the pattern of possible and impossible **coreference**. For example, as pointed out in Moro (2006), an unmodified possessive pronoun in a clearly predicative sentence in English cannot co-refer with the subject. Thus (8a) cannot be interpreted to mean that Omer cooks for himself. To express this, the addition of *own* is obligatory, as in (8b). In contrast, when a possessed noun phrase occurs in an argument position coreference is perfectly possible (8c):

(8) a. Omer$_i$ is his$_{*i}$ cook.
b. Omer$_i$ is his$_i$ own cook.
c. Omer$_i$ met his$_i$ (own) cook.

Moro took this as evidence against the existence of equative sentences *tout court*. But if the obviation effect in (8a) is diagnostic for predicative status, then the following examples have to be given some different analysis since here there is no such effect:

(9) a. Omer's cook produces delicious food. But unfortunately, today Omer himself is cooking for us. And as you can tell, Omer$_i$ is not his$_i$ cook!
b. For a moment there in the bad light I thought Omer$_i$ was his$_i$ cook.

Schlenker (2003) takes examples like (9b) to be "standard identity sentences" but fails to observe that they are strongly asymmetric in interpretation, which is completely unexpected for an equative.

Strikingly, the same kind of contrast in interpretation linked to patterns of co-reference is found even with *become*. That is, *become* allows coreference of an unmodified possessive pronoun with the subject under only one reading; the addition of *own* allows the additional predicative reading:

(10) The magician uttered "Abracadabra!" and Omer$_i$ suddenly became his$_i$ (own) cook.

Just like (9a), without *own* this example loses the interpretation that Omer suddenly started cooking for himself, but it is grammatical under a reading where there is a distinct individual named by *his cook* into whom Omer is transformed. If the impossibility of co-reference is tied to predicative status for the postcopular nominal in some way, we are then forced to conclude that that there are asymmetric (and therefore not equative) copular sentences—including ones with the "semi-copula" *become*—where both of the noun phrases flanking the copula are referring expressions. What is the syntax that underlies this case?

Turning now to a different contrast: one highly salient distinction between predicational and specificational sentences in some languages is the pattern of **agreement**. Within and across languages, agreement in clearly predicational clauses is with the referential/subject noun phrase, but agreement patterns are known to vary in specificational clauses. Some languages (e.g. English) systematically show agreement with the first, structurally higher noun phrase, some (e.g. Italian, Spanish) with the second, and our work shows that others (e.g. Icelandic, Dutch) show variation between those two options. (11b) is from an important early work on this difference (Moro, 1997).

(11) a. The cause of the riot {was/*were you}.

b. La causa della rivolta *è/ sono io. ITALIAN
the cause of.the riot *be.3.SG/ be.1.SG I
'The cause of the riot is me.

c. . . . hvort aðalvandamálið %**væri**/ %**væruð**/ %**væru** þið.
if main problem.DEF be.SBJ.3.SG/ be.SBJ.2.PL/ be.SBJ.3.PL you.PL
'. . . whether the main problem is you.PL' ICELANDIC

The agreement properties of other types of copular clauses are only recently beginning to be investigated (Béjar and Kahnemuyipour 2017, Coon and Keine 2020); further work in this area holds out the prospect of a better understanding of the structures involved, as well as shedding light on the agreement mechanism itself.

Agreement in copular sentences cannot be studied without also investigating **case**. For example, in Dutch "hypothetical" examples like (12a) the second nominal is accusative, while in a specificational sentence like (12b) it is nominative. Different patterns are found even in closely related Germanic languages (Sigurðsson, 2006), but again this has not been explored in any detail.

(12) a. Als ik **jou** was zou ik meteen weggaan. DUTCH
if I.NOM 2SG.**ACC** be.PST.SG would I immediately leave
'If I were you I would leave immediately.'

b. Het enige slachtoffer hierin bent **jij**.
the only victim herein be.PRES.2SG 2SG.**NOM**
'The only victim in this is you.'

Further, there are a number of languages in which equatives are said to require a **pronominal** element in their structure (in addition to an (optional) copular verb) as e.g. in Russian, where the pronominal element *éto* is obligatory in equative sentences. In Hebrew such an element can appear in both predicational copular clauses on the one hand, and equative sentences and specificational sentences on the other, but is obligatory only in the latter types (Doron, 1986, Greenberg, 2003). The nature of these elements and pronominal copulas more generally is disputed and may well differ between languages; as elsewhere, what is taken to be an "identity sentence" is rarely carefully defined, and this is not a trivial matter, neither linguistically nor in philosophy. In Hebrew, for example, the pronominal element has been argued to be a reflex of inflection (Rothstein, 2001, 238); in Russian it has been argued to be an information-structural reflex (Geist and Blaszczak, 2000); for a recent analysis of *ce* in French copular sentences as a reflex of syntactic constraints

related to information structure, see Roy and Shlonsky (2019). The distribution of such pronominal elements within languages where they are variable calls for a better understanding of their syntax.

(13) a. Ciceron - éto byl Tullij — RUSSIAN
Cicero.NOM this.N was.M Tully.NOM
'Cicero was Tully.' (Geist, 2007, 90)

b. Dani *(hu) mar yosef — HEBREW
Dani PRON.M.SG Mr Yosef
'Dany is Mr Yosef.' (Rothstein, 2001, 207)

c. L'étoile du matin *(c')est l'étoile du soir. — FRENCH
the.star of.the morning PRONCE.is the.star of.the evening
'The morning star is the evening star.' (Roy and Shlonsky, 2019, 160)

Finally, some types of copular clauses have distinct **information-structural** properties. One repeatedly noted pattern comes from English specificational copular clauses, where the post-copular noun-phrase is necessarily focused (14B") vs. (15B") , a requirement that is absent in predicational structures (14B') and (15B') (examples from Heycock and Kroch 2002, 148f).

(14) A: Who was the culprit? (John or Bill?)
B': JOHN was the culprit.
B": The culprit was JOHN.

(15) A: What was John? (Was John the culprit or the victim?)
B': John was the CULPRIT.
B": *The CULPRIT was John.

A good empirical and theoretical understanding of these information-structural properties is important for the syntactic analysis and vice-versa. For example, while the syntactic analysis in Mikkelsen (2005), Milway (2020) relies crucially on analysing the first nominal as a topic, Hartmann (2019) takes it rather to be focus that gives rise to inversion. Also considering focus a crucial property, Shlonsky and Rizzi (2018), Roy and Shlonsky (2019) propose movement of the second nominal to a low focus position.

1.4 Internal structure of nominals in copular clauses

Given the widely adopted hypothesis from Partee (1986) that the locus for differences in interpretation between different types of nominal copular clause lies within the nominals themselves, an account of the structure of these nominals is called for. In her seminal work on the interpretation of nominals in English, Partee argued for an invariant semantics for the copula, but that noun phrases could be shifted in semantic type, giving rise to the observed distinctions. While Partee generally did not pursue issues of syntactic encoding, some syntacticians have since developed proposals that tie the different possible interpretations of nominals in copular clauses to differences in their syntactic structure. Notable proposals in this area include Zamparelli 2000 and Roy 2013.

As discussed above, Roy argues for a distinction between two types of nominal **predication**—characterizing and defining—and ties this to different functional heads (NumP and ClassifierP respectively). However, she sets aside definite nominals almost entirely, on the grounds—not seriously defended in her book—that definites do not form predicates. Zamparelli on the other hand does propose that there are projections within the nominal spine that can function as predicates, including when a definite determiner is present (see also Hartmann 2008 for a similar point based on existential sentences). In more recent work (Cheng et al., 2017) this view is elaborated to relate the predicational use of definites to the class of "weak" (non-anaphoric) definites (Schwarz, 2009), and initial steps are taken to incorporate data from classifier languages, in particular Cantonese and Mandarin Chinese. There are however dissenting voices: Julien (2006) in particular has argued that there is no evidence for a syntactic distinction between nominals used in predicate vs. argumental positions. The internal structure of nominal predicates—in particular, definite nominal predicates—therefore remains very much in question.

Even more disputed is the internal structure of nominals involved in **specificational** copular clauses. On the one hand, there are proposals that the *second*, focused, nominal in these cases is the only non-elided part of a larger syntactic structure, akin to a "fragment answer" (Schlenker, 2003), as schematised in (16a); this extends a similar proposal for specificational pseudoclefts elaborated in den Dikken et al. (2000), as schematized in (16b):

(16) a. Her worry is ~~her worry~~ herself.
b. What she saw in the mirror was ~~she saw~~ herself.

The motivation for this proposal is to capture the "connectivity effects" found in specificational sentences: here the possibility of a reflexive in a syntactic structure which if taken at face value—without the proposed hidden structure—would not contain a licensing antecedent in an accessible position. This however requires treating not only deverbal nouns like *worry* as dyadic, but also nouns such as *chair* and *desk*; it also requires a theory of why the ellipsis in cases like (16a) is obligatory. It is a limitation here that the evidence base for "connectivity effects" in these non-cleft copular sentences is much narrower than in pseudoclefts, the cases discussed in the literature generally being limited to the distribution of reflexives (as above) and proper names. As yet there is no systematic investigation of comparable cases in languages with a richer set of anaphors, or of the implications of more recent work on binding/obviation (e.g. Johnson 2019, Charnavel 2020).

In the type of analysis just outlined, special properties of specificational clauses are attributed to unpronounced structure embedding the second nominal. On the other hand, recent proposals due to den Dikken (2006) and taken up in Béjar and Kahnemuyipour (2017, 2018) instead single out the *first* nominal as part of a complex unpronounced structure. In these proposals the first nominal is embedded in something like a silent free relative structure along the lines of (17).

(17) [~~who is~~ the best candidate] is you.

For den Dikken, this structure is important as it forces the inversion referred to above; for Béjar and Kahnemuyipour the crucial aspect of this structure is that the deep embedding of the overt nominal within a clausal structure can explain why in some languages agreement seems to "skip" this nominal and agree with the second, focused nominal, as discussed with respect to (11b) above. It is equally important for this view that the second DP *you* in (17) not be deeply embedded, given that the verb succeeds in agreeing with it. Thus the two sets of proposals for hidden structure are incompatible. As observed above, connectivity effects in specificational sentences have mostly been explored only in English (but see Hartmann et al. 2013 on Hungarian), which does not exhibit the same type of agreement as Italian or German. What is lacking so far is a systematic investigation of connectivity effects in the latter type of languages; this would enable us to determine whether the label of specificational clause has actually been applied to very different structures, depending on the language, or whether in fact we must reject one or both of these approaches advocating hidden clausal structure.

As for the internal structure of the nominals in **equative** clauses, there is very little discussion to be found in the literature. On the one hand, in some of the recent literature it is claimed that equatives simply don't exist, that all putative cases can be reduced to predication or specification interpreted as "inverted predicaton" (see e.g. Moro 2006). Elsewhere it is sometimes assumed that little needs to be said, given the assumption that both nominals are referring expressions and that there is a distinct "equative copula" that effectively functions as a transitive verb. Den Dikken's work stands out in proposing a complex asymmetric structure for equative clauses like *Cicero is Tully* that is based on proposals originally made for Scottish Gaelic in Adger and Ramchand (2003). It is striking however that despite the title of that paper—*Predication and Equation*—the authors show that the prototypical "Cicero/Tully" case of equation that den Dikken is analysing in English is simply ungrammatical in Scottish Gaelic; the examples that they cite all involve roles (*the teacher* or *Hamlet*) and are not grammatical if the two nominals are switched.

1.5 Cross-linguistic studies

While there are a number of studies on copular clauses in individual languages (e.g. the chapters in Arche et al. (2019)), there are only a few studies that look at nominal copular clauses considering a range of languages and a number of different predicate sentences. In the typological literature on copular clauses the main focus is on the nature of the copular element and in which contexts it can be dropped (see Stassen 1997, Pustet 2003, Stassen 2013). Copular elements can either be verbal, pronominal (including demonstratives) or particle-like; they can be dropped in different environments (present tense, 3rd person) or be available only with a subset of predicative projections (adjectival, nominal or verbal projections). Even more than in the theoretical literature above there is hardly any in-depth analysis of different types of copular sentences with definite "predicates". For example, the cross-linguistic investigation in Pustet (2003) largely sets aside any copular constructions where the second nominal "conveys uniqueness" (pp. 30–31), and does not distinguish between predicative and referential uses of definites.

1.6 Open research questions to be addressed

In summary, the following are research questions that arise from the current state of the literature and that we plan to address in our work:

- What are the main types of nominal copular clauses that can be distinguished on interpretive grounds for which we also have evidence in a wide range of languages of syntactic differentiation? (Work package 1)
- What diagnostic tools can be developed in order to allow for systematic cross-linguistic comparison of the properties of these types? (Work packages 1,4)
- What is the nature of the syntactic structures in which nominals in copular clauses are embedded? How can their varying case, agreement properties and information structure—varying both within and between languages—be accounted for? What explains their differing patterns of coreference? What determines the distribution within each language of a copular verb and/or pronominal element? (Work package 2)
- What is the internal syntactic structure of the nominals in copular clauses, and what can we learn from that about the internal structure of nominals more generally? Do different types of predication involving definite "predicates" follow from the presence of different functional layers within definite nominals that can be identified across languages? (Work package 3)
- What, if any, connectivity effects obtain in non-cleft copular sentences? Is there evidence for ellipsis or other silent structure? Or do apparent connectivity effects instead force us to rethink the syntactic/semantic/pragmatic character of e.g. binding conditions and obviation effects? (Work packages 2,3)

research that we have done that serves as groundwork on these questions; In subsequent sections we give more detail on our approach, working hypotheses, and methodology.

2 Objectives and work programme

2.1 Anticipated total duration of the project

The project is planned for 36 months.

2.2 Objectives

The summary of the current state of the art above has indicated a number of open questions, both empirical and theoretical, that relate to overarching theoretical issues, particularly concerning the

mapping between syntax on the one hand and information structure and interpretation on the other. This project aims to address these by setting the following specific objectives:

1. Establish the range of subtypes of nominal copular clauses for which there is evidence for both semantic and syntactic differentiation. Develop well-defined, replicable semantic and syntactic criteria to enable systematic crosslinguistic comparison.
2. Develop detailed analyses of the syntactic structures embedding nominal predicates that are associated with the different subtypes of nominal copular clauses in a selected set of different languages, including languages that are typologically and genetically unrelated. Establish the information-structural properties specific to particular subtypes, and the relation between syntactic structure and information-structural properties.
3. On the basis of in-depth analysis of nominal copular clauses in individual languages, combined with detailed crosslinguistic comparison, show how the internal structure of nominals, including in particular definite nominals, contributes to and constrains their possible predicative uses. Determine to what extent ellipsis is implicated in cases showing unexpected patterns of coreference.
4. Develop and make freely available templates for further comparative research that are designed for use by the wider research community and will be advertised to the community as an Open Science resource.

Below we set out the work packages: project members are abbreviated as JH (Jutta Hartmann), CH (Caroline Heycock), VH (Veronika Hegedűs), IR (Isabelle Roy), RZ (Roberto Zamperelli), PD-Bi (Bielefeld-based postdoc), PD-Ed (Edinburgh-based postdoc).

2.3 Work programme including proposed research methods

In order to achieve our objectives, we will build on our previous work on nominal copular clauses, extending it in three dimensions. First, going beyond agreement, we will investigate in detail a wider range of grammatical phenomena (including case, binding, word order, embedding possibilities) and information-structural effects (see WP 2a,b). Second, we will build on CH's collaboration with RZ, and IR's work on nominals, to develop an analysis of the syntax and semantics of the internal structure of predicative nominals, focusing in particular on definites, and aiming for a consistent analysis that extends to connectivity effects (WP3). Third, the project will extend our work on Germanic to other typologically different languages to broaden the empirical basis for our theoretical proposals concerning all the aspects just mentioned. One language we will focus on is Hungarian, in collaboration with the Mercator fellow Hegedűs (VH). Hungarian has a rich agreement system and, in contrast to the Germanic languages, an overt encoding of information-structure. IR and RZ bring expertise in two different Romance varieties (French and Italian); IR also has an ongoing collaboration relating to French and to Hebrew with Shlonsky, who will be invited to two of our project meetings, and in addition much of IR's recent research focus, carried out with members of her lab in Nantes, is on extending detailed work on nominal copular constructions to a wider range of languages, including the Niger-Congo language Wolof and the Western Mande language Bambara. Further, the two postdocs will be selected with an eye to adding to the typological breadth of the project by bringing in expertise in crosslinguistic work, and knowledge of one or more languages most relevant to studying aspects of the internal structure of the DP (PD-Ed) and aspects of the sentential syntax (i.e. agreement, case, word order, embedding, information-structure) of copular clauses (PD-Bi).

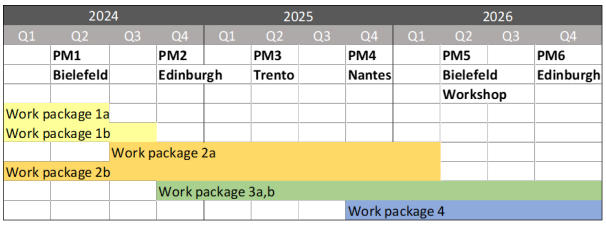

Figure 1: Proposed timeline

Our research will be organised into four work packages, with the timeline indicated in Figure 1 (PM=project meeting);

Work package 1: In this work package we extend the current understanding of different types of nominal copular clauses in order to investigate their grammatical properties. Following Higgins' work and subsequent literature, we start out with the four-way classification of predicational, specificational, identificational and equative copular clauses. For predicational clauses, we will investigate whether there is any basis for extending Roy's subclassification of indefinite predicate nominals to definites. We will also undertake a thorough review of cases where both nominals appear to be referring expressions (for example, both positions can be proper names or pronouns). While this is typically assumed to be the hallmark of equatives, we will consider also naming constructions (*I am Samira*), and a range of cases that we referred to above as "counterfactual equatives". This latter group includes conditionals (*If I were you*), dream and role-playing contexts (*Sean Connery is James Bond*), and mistaken identity (*in the darkness she thought I was you*). These cases fail most of the commonly used diagnostics for the predicational type, but they also differ clearly from "typical" equatives in that they are asymmetric; switching the noun phrases gives rise to an entirely different interpretation (*Cicero is Tully* $=$ *Tully is Cicero* vs.*They thought I was you* $\neq$ *They thought you were me*).

In order to investigate to what extent the observed interpretive differences have any syntactic basis, we will develop scenarios to clearly establish the semantic distinctions, and sharpen known syntactic criteria (including binding, case, agreement, inversion, information-structural effects, embedding), extending these also to the less-examined cases mentioned above (**WP1a**). To provide a broader cross-linguistic perspective, we will survey nominal copular clauses in a number of typologically distinct languages (**WP1b**), potentially extending the list of grammatical properties that serve as criteria for distinguishing subtypes.

In **WP1a**, we will (i) conduct a literature review to bring together existing criteria and diagnostics for subtypes of nominal copular clauses; (ii) supplement categories from the theoretical literature with any additional distinctions motivated by the findings from work package 1b; (iii) critically evaluate and systematise diagnostics for the different categorisations. In **WP1b**, we will survey a range of genetically different languages, namely a sample of about 20 languages selected from the WALS online 200 language sample. Our sample will take two to four languages per larger area (i.e. Africa, Australia, Eurasia, North America, South America, Papunesia) each from a different genus (see Dryer 1989 for the relevance of genera for language sampling). The specific language sample will be determined on the basis of availability of reliable sources as well as breadth of variation in the already known relevant factors (presence or absence of copula, languages with definite determiners vs. those without, different types of copular clauses, different agreement and case marking systems, expression of information structure). WP1a will be worked on by CH,JH,RZ,IR,VH,PD-Bi; WP1b will be mostly conducted by PD-Bi and JH, with input from both IR and VH.

The resulting map of types of copular clauses and grammatical properties is the basis both for our in-depth analysis of the different types in a subset of languages, and for the template for further cross-linguistic investigation (WP4) to underpin further research on copular clauses.

Work package 2: Based on these distinctions, we turn to the detailed analysis of the syntactic properties of nominal copular clauses in a chosen set of individual languages. Based on the types isolated in WP1, we intend to explain the relation between specific readings and their grammatical properties. The strongest structural hypothesis is that each type is associated with a different syntactic structure, whereas the strongest semantic hypothesis is that these differences result from semantic type shifting alone. Based on our previous work on specificational copular clauses it is evident that this strongest semantic hypothesis is not adequate, at least for those languages that we have studied. Thus, our aim here is to investigate to what extent the strongest structural hypothesis is adequate for other types of copular clauses and for a range of different structures. In **WP2a** we concentrate on the grammatical properties such as case, agreement and embedding; we include the dimension of information-structure in **WP2b**, see below. We will start from the assumption that predication is a structural configuration, subjecting current proposals such as

PredP (Bowers, 1993), RelP (den Dikken 2006) and Moro's dynamic antisymmetry (Moro, 2000) to closer scrutiny and considering recent counterproposals, e.g. Matushansky (2019) and alternative proposals that a small clause is an extended projection of the predicate (aP,nP,pP) in parallel to the analysis of prepositional phrases and particles (Hegedűs, 2013).

Taking the inversion analysis of specificational copular clauses as one source of the varied agreement patterns that we have already documented in this type of copular clause in Germanic, we will extend our work to other types, in particular equatives (including what were termed above "counterfactual equatives"), investigating the possibility of multiple agreement (Coon and Keine, 2020) as another source of such variation. Besides the effects of agreement, case is another important grammatical feature that can indicate a complex structure of the DP (e.g. due to ellipsis or additional functional structure blocking case relations), or even argumental status. The availability of embedding in non-finite structures is a potential indication of whether inversion is implicated for the given type. Last but not least the conditions under which the copula *be* and other copular verbs appear, when they can or must be absent, indicates both the minimal structure for the availability of a specific reading as well as the nature and position of the copula itself.

In **WP2b** we will investigate the interaction of the syntax of nominal copular clauses with constraints on information structure. As noted above, in previous work it has been shown that specificational copular clauses require a specific information structure (see (14),(15) above); we believe the correct characterization to be that the second DP needs to be focused (see Heggie 1988, Heycock 1994b, Heycock and Kroch 2002, for experimental evidence see Hartmann 2019). No such restriction holds for predicational copular clauses. This seems to be a defining feature of specificational copular clauses not only in English, but also in other languages (see e.g. Hartmann et al. 2013). To what extent equative sentences may be associated with particular information structural configurations is not as clear. Information-structural asymmetries might be responsible for some of the asymmetries in equatives, without other syntactic differences, so it is important to consider these differences as well. We will look into this question starting on the basis of Hungarian, where information structure shows a well-studied syntactic encoding, extending our work to other languages that are uncovered in WP1.

Our work in WP2 will be based on data from language consultants (questionnaires, field work in the project locations as well as in the language communities of the additional language(s) in the expertise of PD-Bi) and IR and her lab, together with experimental studies. We have found in our previous work that setting up comparable experimental studies across languages is important particularly where work with consultants reveals differences between languages in the degree of inter-speaker and/or intra-speaker variation, for example with respect to agreement and case. We are planning for two experiments per language in this WP, extending our work on agreement, but including further factors relevant for the types of copular clauses in the languages under investigation isolated in WP1. Work package 2 will be the main responsibility of the Bielefeld team, and the major domain of research for PD-Bi; the work will be done in close collaboration with the Edinburgh team to include work on the languages worked on within IR's lab, and to have an eye on considerations of DP-internal structure throughout.

Work package 3: One of the most radical proposals for the simplification of the taxonomy of copular clauses is that of Moro (2006), who proposes that all putative equative cases can be reanalysed as predicational or specificational, and that, in turn, specificational copular clauses are effectively a derivational variant of predicational ones. As indicated earlier, however, we believe that there are reasons—both theoretical and empirical—to question this approach, and that crucial aspects of the differentiation are likely to derive from the internal syntax of the nominals involved.

While there have been a number of different proposals concerning the possibility that indefinite nominal predicates may involve functional projections smaller than a "full" DP (Roy 2013 is just one example), there has been much less attention to the possibility of "reduced" definite predicates. In **WP3a** we will take further the proposals outlined in Cheng et al. (2017) to develop a theory

of the structural contribution to the predicative use(s) of definites, relating this also to existing proposals for indefinites. CH will work with PD-Ed, RZ and IR to test the hypothesis that there are distinct functional projections within the DP that can function as predicates, the highest of which also corresponds to the "weak definite" arguments of Schwarz (2009) and which hosts the English definite determiner *the*. A crucial auxiliary hypothesis that will also be tested is the MINIMIZE STRUCTURE principle of Cheng et al. (2017): namely that language uses the smallest category that can accommodate all the material in the numeration and that has or can be converted to the necessary semantic type. This analytical work will draw on Germanic and Romance but also include comparative work on at least one language where definiteness is not associated with an overt definite article; we will recruit the Edinburgh-based research assistant (PD-Ed) to bring research expertise in relevant languages. In this work package we will also test the hypothesis that equatives of all subtypes involve additional structure above the level of the "strong definites." The empirical basis for this part of the research will draw importantly on the crosslinguistic work in WP1b, since we expect it to yield new evidence concerning the crosslinguistic distribution of equatives, something which currently is very hard to assess given the widely different assumptions about the definition of this class across the literature.

In the second strand of this work package on the internal structure of the nominals involved in nominal copular clauses, **WP3b**, we will systematically investigate the extent to which "connectivity effects"—in particular the unexpected licensing of reflexives and other anaphors, related obviation effects, and bound readings of pronouns—appear in specificational clauses other than pseudoclefts (although comparisons will also be drawn with the already much more thoroughly investigated phenomena in pseudoclefts). Since coreference and binding interact with information structure in ways that can create confounds (Lahousse, 2009) this work package will draw on the work in WP2b. CH will work with PD-Ed, and draw on the expertise of IR and RZ as well as the Bielefeld team to construct test cases both in languages that exhibit agreement with the first nominal (e.g. English, French) and those that exhibit agreement with the second (e.g. Italian, German). We will incorporate the insights of recent work on anaphora and obviation (e.g. Johnson 2019, Charnavel 2020) and will include in the languages studied in detail those which have a richer inventory of reflexive forms (e.g. Icelandic). In order to ensure the robustness of the data we anticipate supplementing work with individual language consultants with experimentally controlled questionnaire studies in five selected languages. We will use our data to test the hypothesis that there is hidden clausal structure in either the first (den Dikken, 2006) or the second (Schlenker, 2003) nominal in such copular clauses. In turn we will explore the consequences of the results for the formulation of binding and locality principles.

Work package 4: Apart from the open theoretical questions, it is striking that there is hardly any systematic cross-linguistic work on agreement, case or word order properties of different types of copular clauses (see section 1.2 and 1.5) above. Based on our theoretical and empirical insights in WP1-3, we want to facilitate the investigation of copular clauses particularly in lesser studied languages, by developing a questionnaire that will be made freely available, to facilitate research by the wider community on the grammatical properties of nominal copular clauses of different types. Once developed, this questionnaire will then be implemented in an online database such as TerraLing to allow for long-term storage and accessibility of the data from a larger set of languages. TerraLing (`https://www.terraling.com`) is a platform that brings together experts on a specific linguistic phenomenon with experts on specific languages. We will provide a template on the grammatical properties of different types of copular clauses with TerraLing, allowing language experts to provide their expertise. This template will be presented and discussed at the workshop in year 3 and we will be encouraging its future use by the participants at that workshop as well as advertising it to the wider community. We will invite the participants at the Year 3 workshop and also selected additional researchers on lesser studied languages to work with the template and publish their results in an edited volume with a collection of papers on copular clauses, but the longer-term intention is that the template/questionnaire will be an Open Science resource available

to the entire community.

3 Bibliography

Adger, D. and G. Ramchand (2003). Predication and equation. *Linguistic Inquiry 34*(3), 325–359.

Arche, M. J., A. Fábregas, and R. Marín (Eds.) (2019). *The Grammar of Copulas across Languages*. Oxford Studies in Theoretical Linguistics. Oxford, New York: Oxford University Press.

Béjar, S. and A. Kahnemuyipour (2017). Non-canonical agreement in copular sentences. *Journal of Linguistics 53*(3), 463–499.

Béjar, S. and A. Kahnemuyipour (2018). Not all phi-features are created equal: A reply to Hartmann and Heycock. *Journal of Linguistics 54*(3), 629–635.

Bowers, J. (1993). The syntax of predication. *Linguistic Inquiry 24*(4), 591–656.

Charnavel, I. (2020). *Locality and logophoricity: A theory of exempt anaphora*. Oxford: OUP.

Cheng, L. L.-S., C. Heycock, and R. Zamparelli (2017). Two levels for definiteness. In M. Y. Erlewine (Ed.), *Proceedings of GLOW in Asia XI - Volume I*, Volume 84 of *MIT Working Papers in Linguistics*.

Coon, J. and S. Keine (2020). Feature gluttony. *Linguistic Inquiry 0*(0), 1–56.

Coppock, E. and D. Beaver (2015). Definiteness and determinacy. *Linguistics and Philosophy 38*, 377–435.

Declerck, R. (1988). *Studies on Copular Sentences, Clefts and Pseudo-clefts*. Dordrecht: Leuven University Press.

den Dikken, M. (2006). *Relators and Linkers: The Syntax of Predication, Predicate Inversion and Copulas*, Volume 47 of *Linguistic Inquiry Monographs*. Cambridge, MA: MIT Press.

den Dikken, M. (2017). Pseudoclefts and other specificational copular sentences. In M. Everaert and H. C. van Riemsdijk (Eds.), *The Wiley Blackwell Companion to Syntax* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

den Dikken, M., A. Meinunger, and C. Wilder (2000). Pseudoclefts and ellipsis. *Studia Linguistica 54*(1), 41–89.

Doron, E. (1986). The pronominal "copula" as agreement clitic. In H. Borer (Ed.), *The Syntax of Pronominal Clitics*, pp. 313–332. New York: Academic Press.

Dryer, M. S. (1989). Large linguistic areas and language sampling. *Studies in Language 13*(2), 257–292.

Geist, L. (2007). Predication and equation in copular sentences: Russian vs. English. In I. Comorovski and K. von Heusinger (Eds.), *Existence: Semantics and Syntax*, Volume 84 of *Studies in Linguistics and Philosophy*, pp. 79–105. Dordrecht: Springer.

Geist, L. and J. Blaszczak (2000). Kopulasätze mit den pronominalen Elementen to/èto im Polnischen und Russischen. In E. Lang (Ed.), *ZAS Papers in Linguistics*, Volume 16, pp. 115–139. Berlin: ZAS.

Greenberg, Y. (2003). *Manifestations of Genericity*. London: Routledge.

Hartmann, J. M. (2008). *Expletives in Existentials: English there and German da*. Utrecht: LOT.

Hartmann, J. M. (2018). A focus analysis of apparent predicational clefts. *Belgian Journal of Linguistics 32*(1), 121–143.

Hartmann, J. M. (2019). Focus and prosody in nominal copular clauses. In S. Featherston, R. Hörnig, S. von Wietersheim, and S. Winkler (Eds.), *Experiments in Focus: Information Structure and Semantic Processing*, pp. 71–103. Berlin: De Gruyter Mouton.

Hartmann, J. M., V. Hegedűs, and B. Surányi (2013). Pseudoclefts in Hungarian. In J. Brandtler, V. Molnár, and C. Platzack (Eds.), *Papers from the 2011 Lund conference*, Volume 13 of *Approaches to Hungarian*, pp. 67–96. Amsterdam: John Benjamins.

Hartmann, J. M. and C. Heycock (2017). Variation in copular agreement in Insular Scandinavian. In H. Thráinsson, C. Heycock, H. P. Petersen, and Z. S. Hansen (Eds.), *Syntactic Variation in Insular Scandinavian*, Volume 1 of *Studies in Germanic Linguistics (SIGL)*, pp. 233–275. Amsterdam: John Benjamins.

Hartmann, J. M. and C. Heycock (2019). (Morpho)syntactic variation in agreement: Specificational copular clauses across Germanic. *Frontiers in Psychology, Language Sciences 10*(2994).

Hartmann, J. M. and C. Heycock (2022, June). Person effects in agreement with Icelandic low nominatives: An experimental investigation. *Natural Language and Linguistic Theory*.

Hegedűs, V. (2013). *Non-Verbal Predicates and Predicate Movement in Hungarian*. Utrecht: LOT.

Heggie, L. (1988). *The Syntax of Copular Structures*. Ph. D. thesis, Univ. of Southern California.

Heycock, C. (1994a). The internal structure of small clauses. In J. Beckman (Ed.), *Proceedings of NELS 25*, Volume One, Amherst, Mass., pp. 223–238. GLSA.

Heycock, C. (1994b). *Layers of Predication: The Non-Lexical Syntax of Clauses*. New York: Garland.

Heycock, C. (2012). Specification, equation, and agreement in copular sentences. *Canadian Journal of Linguistics 57*(2), 209–240.

Heycock, C. (2013). The syntax of predication. In M. den Dikken (Ed.), *The Cambridge Handbook of Generative Syntax*, pp. 322–352. Cambridge: Cambridge University Press.

Heycock, C. (2021). Copular sentences. In L. Matthewson, C. Meier, H. Rullmann, T. E. Zimmerman, and D. Gutzmann (Eds.), *The Wiley Blackwell Companion to Semantics*. London: Wiley-Blackwell.

Heycock, C. and A. Kroch (1999). Pseudocleft connectedness: Implications for the LF interface level. *Linguistic Inquiry 30*(3), 365–397.

Heycock, C. and A. Kroch (2002). Topic, focus, and syntactic representations. In L. Mikkelsen and C. Potts (Eds.), *Proceedings of WCCFL 21*, Somerville, Mass., pp. 101–125. Cascadilla Press.

Higgins, F. R. (1979). *The Pseudo-Cleft Construction in English*. New York: Garland.

Jacobson, P. (1999). Towards a variable-free semantics. *Linguistics and Philosophy 16*, 353–422.

Johnson, K. (2019). Principle C. Ms., UMass, `https://people.umass.edu/kbj/homepage/Content/PC_final.pdf`.

Julien, M. (2006). Nominal arguments and nominal predicates. In J. M. Hartmann and L. Molnárfi (Eds.), *Comparative Studies in Germanic Syntax: From Afrikaans to Zurich German*, Linguistik Aktuell/Linguistics Today, pp. 115–140. John Benjamins Publishing Company.

Lahousse, K. (2009). Specificational sentences and the influence of information structure on (anti-) connectivity effects. *Journal of Linguistics 45*, 139–166.

Matushansky, O. (2019). Against the PredP theory of small clauses. *Linguistic Inquiry 50*(1), 63–104.

Mikkelsen, L. (2005). *Copular Clauses: Specification, Predication and Equation*, Volume 85 of *Linguistik Aktuell*. Amsterdam: John Benjamins.

Milway, D. (2020). The contrastive topic requirement on specificational subjects. *Canadian Journal of Linguistics/Revue canadienne de linguistique 65*(2), 181–215.

Moro, A. (1997). *The Raising of Predicates: Predicative Noun Phrases and the Theory of Clause Structure*. Cambridge: Cambridge University Press.

Moro, A. (2000). *Dynamic Antisymmetry*. Cambridge, MA: MIT Press.

Moro, A. (2006). Copular sentences. In M. Everaert and H. van Riemsdijk (Eds.), *The Blackwell Companion to Syntax*, Volume II, Chapter 18, pp. 1–23. Oxford: Blackwell.

Partee, B. (1986). Ambiguous pseudoclefts with unambiguous *be*. In S. Berman, J.-W. Choe, and J. McDonough (Eds.), *Proceedings of NELS 16*, Amherst, MA, pp. 354–366. GLSA.

Pustet, R. (2003). *Copulas: Universals in the Categorization of the Lexicon*. Oxford Studies in Typology and Linguistic Theory. Oxford: Oxford University Press.

Rothstein, S. (2001). *Predicates and their Subjects*. Dordrecht, Boston: Kluwer.

Roy, I. (2013). *Nonverbal Predication: Copular Sentences at the Syntax-Semantics Interface*. Oxford: Oxford University Press.

Roy, I. and U. Shlonsky (2019). Aspects of the syntax of *ce* in French copular sentences. See Arche et al. (2019), Chapter 8.

Russell, B. (1905). On denoting. *Mind 11*, 479–493.

Schlenker, P. (2003). Clausal equations (a note on the connectivity problem). *Natural Language*

and Linguistic Theory 21, 157–214.

Schwarz, F. (2009). *Two Types of Definites in Natural Language*. Ph. D. thesis, UMass, Amherst.

Shlonsky, U. and L. Rizzi (2018). Criterial freezing in small clauses and the cartography of copular constructions. In J. M. Hartmann, M. Jäger, A. Kehl, A. Konietzko, and S. Winkler (Eds.), *Freezing: Theoretical Approaches and Empirical Domains*, pp. 29–65. De Gruyter Mouton.

Sigurðsson, H. Á. (2006). The Nom/Acc alternation in Germanic. In J. M. Hartmann and L. Molnárfi (Eds.), *Comparative Studies in Germanic Syntax*, Volume 97 of *Linguistik Aktuell/Linguistics Today*, pp. 13–50. Amsterdam: Benjamins.

Stassen, L. (1997). *Intransitive Predication*. Oxford: Clarendon Press.

Stassen, L. (2013). Zero copula for predicate nominals. In M. S. Dryer and M. Haspelmath (Eds.), *WALS Online*.

Zamparelli, R. (2000). *Layers in the Determiner Phrase*. New York: Garland.