CHAPTER 2

What is nominalization? Towards the theoretical foundations of nominalization

Masayoshi Shibatani
Rice University

過而不改，是謂過矣
(孔子)

This paper discusses foundational issues in nominalization, focusing on empirical, conceptual, and theoretical problems that have plagued the field for years. Current definitions of nominalization are based on narrow observations on verbal-based nominalizations, while languages across the globe display nominal-based ones, many of which share morphology with the former. Nominalization morphology in many languages also applies to units larger than words, yielding grammatical nominalizations besides lexical nominalizations. The imbalance in the past studies, which have focused on so-called relative clauses at the expense of grammatical nominalizations, has resulted in a lopsided view on the relationship between the two. This, in turn, has led to the mishandling of different manifestations of nominalizations as if they are derivatives of relative clauses, as suggested by the widely used terms such as "headless relative clause" and "internally-headed relative clause". We demonstrate that these, including the ordinary restrictive relative clause, are not independent grammatical structures but are epiphenomena arising from the uses of grammatical nominalizations. A clear distinction between structures and their use is a theoretical prerequisite in arriving at a satisfactory understanding of the nature of grammatical nominalizations and their role in grammar.

1. Introduction

Studies on nominalization, in both Western and Eastern grammatical traditions, have largely concentrated on lexical nominalizations, neglecting grammatical nominalizations, despite their theoretical importance and far-reaching implications to the descriptive practice. This imbalance is due to the fact that while lexical nominalizations (e.g. English *sing-er*) typically involve distinct morphology and
their lexical status as nouns is relatively clear-cut, grammatical nominalizations (e.g. [I know] that John recklessly shoots trespassers; [I saw] John shoot trespassers; John’s recklessly shooting trespassers [angered the entire community]; To shoot trespassers [is unacceptable]) vary considerably in form, some of which displaying internal structural properties similar to clauses, and their nominal status is less fully realized compared to lexical nominalizations (e.g. a/the shooting [of trespassers]; those terrible shootings [of trespassers], but not *a/*the shooting trespassers [is unacceptable]).

Our main concerns in this paper are these understudied types of nominalization, whose position in grammar, we claim, has not been fully assessed despite their far-reaching theoretical implications. Our emphasis is on the conceptual and empirical issues pertaining to grammatical nominalizations, because there are deeprooted and widespread misconceptions about them, hampering a proper perspective on and correct understandings of the relevant data. Because of this, we provide a multitude of data, not only from the Americas, but also elsewhere around the globe that bear on the topic, which to a great extent are redundant, but which are felt necessary to dispel the misunderstandings found in many, if not all, current definitions of nominalization. An abundance of data also helps correct some narrow views on nominalization found in several statements and analyses in professional publications, such as “a somewhat more rare function of nominalization [is its use] as a relative clause modifying a head noun” (Comrie & Thompson 2007: 378; emphasis added) and “in certain languages relativization is indistinct from nominalization” (Comrie & Thompson 2007: 379; emphasis added).

This paper is organized as follows. In the next section the current definitions of nominalization are placed under close scrutiny, with our findings that they are either incorrect or at best insufficient. We offer an alternative definition of nominalization as a metonymy-based phenomenon that applies not only to verbals but also to nominals and that yields structures that form nouns as well as phrasal units. Section 3 starts earnest discussions of verbal-based grammatical nominalizations. A major distinction is made between event nominalizations (Section 3.1) and argument nominalizations (Section 3.2) and their formal representations are attempted in Section 4. Section 5 starts discussions of the usage patterns of these grammatical nominalizations, dividing them into two major uses, the NP-use and the modification-use (Section 5.1). The latter includes an adverbial use of nominalizations, which are also used in so-called clause-chaining constructions. Section 5.2 demonstrates that so-called internally-headed relative clauses are not really relative clauses at all and that they are no more than an instance of NP-use

1. See Lees (1963) for an early, but still the most comprehensive treatment of English nominalizations.
of event nominalizations. Section 5.3 discusses another issue of whether so-called complement clauses are distinct from nominalizations in NP-use.

Similar arguments are advanced for argument nominalizations, first showing that so-called headless relative clauses are an NP-use of argument nominalizations and are not relative clauses at all, contrary to the widely subscribed labeling of them as such (Section 5.4). Then, Section 5.5 demonstrates that ordinary relative clauses are no more than a modification-use of argument nominalizations and that there are no structures such as relative clauses independently from argument nominalizations in modification-use. A new analysis of relative clauses is proposed that minimizes the role of so-called relative pronouns that play a significant role in the traditional as well as generative analyses of relative clauses.

In view of the lack of proper understandings of what clauses and sentences are in the field, Section 6.2 offers functional definitions of clauses and sentences and shows how they differ from nominalizations. Section 6.3 discusses the phenomenon of insubordination, by which event nominalizations may become used as sentences. We then go on to provide some evidence supporting our claim that nominalizations are neither clauses nor sentences (Sections 6.4 and 6.5).

Section 7 takes up nominal-based nominalizations showing that many languages of the world show a morphological connection between them and verbal-based nominalizations. The importance of nominal-based nominalizations is shown in Section 8, where it is demonstrated that the NP-use of N-based nominalization is the locus of innovations in the development of NP-use markers, which, by spreading to the modification-use of both N-based and V-based nominalizations, may eventually become nominalizers. Finally, Section 9 takes stock of our lengthy discussions on grammatical nominalizations and demonstrate their implications for both theoretical and descriptive issues inherent in the treatment of relative clauses by Keenan and Comrie (1977) and Comrie and Keenan (1979).

2. Defining nominalization

Popular definitions of nominalization found in dictionaries, encyclopedias, and introductory textbooks, as well as more technical scholarly papers such as the following are all problematic.

nominalization refers to the process of forming a noun from some other word-class or, especially in classical transformational grammar, the derivation of a noun phrase from an underlying clause. (Crystal 1980: 328)

...operations that allow a verb to function as a noun ...are called nominalizations, and can be described with a simple formula: \( V \rightarrow N \). (Payne 1997: 223)
The term *nominalization* means 'turning something into a noun'.
(Comrie & Thompson 1985/2007: 334)

‘nominalization’ actually conflate[s] two properties: deverbalization ...and substantivization (acquisition of noun-properties). (Malchukov 2004: 6)

Below we focus on three issues that these definitions raise, namely (i) the overall characterization of the process, (ii) the inputs to the process, which are restricted to verbs or members of non-nominal categories (Crystal, Payne, Malchukov), and (iii) the outputs, defined as nouns (Payne, Comrie and Thompson, Crystal), or noun phrases (Crystal).

### 2.1 Nominalization as a metonymic process

The definitions of nominalization above are all too general in that they do not specify the relationships between the inputs and the outputs other than that the former are verbal and the latter nominal. It is, however, not the case that any type of nominalization is derivable from any verbal input. Imagine an extreme case of trying to derive forms using the English nominalization suffix -er such as *singer* and *driver* from verb roots like *walk* and *kill*. No language would allow derivations like *walk > singer* and *kill > driver*, while all the popular definitions of nominalization simply say that the process derives nouns or nominal expressions from verbs or members from non-nominal categories. Along a similar but more plausible line, consider deriving nominalizations denoting agents like *singer* and *driver* from stative or simple processes verbs such as *resemble* and *die* that denote non-activity relational properties, e.g. *resembler* and *dier*. Agentive nominalizations are derivable only from activity verbs that predicate over an agentive subject. Simple activity verbs, on the other hand, would not yield resultative nominalizations of the type such as (*a*) *painting*, (*a*) *building*, and (*his*) *writings*, which are associated with verbs of production and certain change-of-state verbs. In other words, nominalization is far more constrained than the overall characterizations of the conventional definitions have it.

In view of the relationships between the inputs and the outputs seen above, we define nominalization as a metonymic process along the line of Fillmore’s frame semantics (Fillmore 1976, 1982).\(^2\) Namely, verbal-based nominalization, for example, yields only those forms that denote things and thing-like entities (both

\(^2\) Allan (2001: 251) characterizes a semantic frame as consisting of "characteristic features, attributes, and functions of a denotatum, and its characteristic interactions with things necessarily or typically associated with it" (emphasis added). A similar effort is seen in Langacker’s (1987) Cognitive Linguistics framework in terms of the notion of "profiling".
concrete and abstract) that are in close association with the scenes/scripts evoked by the use of specific verbs. The resulting nominalization structures often have marking that more narrowly circumscribes the range of denotations they evoke, as in the English forms employment, employer, and employee. As these examples show, nominalization yields structures denoting substantive or entity concepts that are metonymically evoked by the nominalization structures themselves such as events, facts, and propositions (cf. employment, employing), as well as concrete objects such as event participants (agents and patients) and entities conventionally associated with specific events like instruments, resulting objects and locations (cf. employer, employee, plier, a building, a landing). As products, nominalizations are like nouns (hence the term “nominalization”) by virtue of their association with an entity-concept denotation, a property that provides a basis for the referential function of a noun phrase headed by such nominalizations. Verbs and verb phrases, on the other hand, are associated with relational concepts (time-stable or transient properties pertaining to an entity or entities) and play a predication function in a clause by ascribing a relational concept to the referent of a subject noun phrase (see Section 6.2). They differ crucially from nouns and nominalizations in not denoting things and thing-like entity-concepts and thereby in being unable to play a referential function.

Metonymy is a powerful cognitive process that allows a variety of form-concept connections increasing the expressive power of a language with limited resources. By taking advantage of our knowledge that many things in the world occur in close association, a metonymic construal allows us to conceptualize and denote entities in alternative ways beyond the conventional form-concept connections. The result of such a process is a richer (and often colorful) description of a denoted entity focusing on some associated features that the speaker finds more informative and relevant to the context. Specific metonymic expressions are not random but are manifestations of experientially-grounded general conceptual metonymic schemas of the type, the part for the whole (Get your butt over here!), producer for product (He bought a Ford), the place for the institution (The White House isn’t saying anything), etc. (Lakoff & Johnson 2003: 38). Schemas that play important roles in nominalization include the event for the state/process/activity, the event for the fact, the event for the proposition (all for event nominalizations), the event for the protagonists

3. A good description of nominalization intimates metonymic form-meaning relationships involved in nominalizations; e.g., “A nominalization on an oblique (i.e., non-subject) argument … typically denotes an object or location closely associated with generic performances of the activity named by the verb from which it is derived.” (Miller 2001: 120; emphasis added).

4. See Section 4 on the difference between denotation and reference.
(argument nominalizations), the event for the result (resultative nominalizations), possessor for possessed, producer for product (both for nominal-based nominalizations), and others.

A single metonymic expression may denote a variety of entity concepts that are closely associated with the concepts denoted by the original words or larger structures, and it is the speech context that determines and selects the denotation/reference most relevant to the context per Gricean maxims of conversation (or the Cooperative Principle), one of which (the Maxim of Relevance) requires an expression to be contextually relevant at the time of the utterance. For example, the United States may metonymically evoke and denote a variety of entities closely associated with the country by this name (e.g. different types of representatives of the country in question), but only a contextually relevant interpretation would be intended by the speaker and would be chosen by the hearer – e.g. the sitting US president in the United States has decided to pull out from the Paris agreement, or a US women’s soccer team in the United States defeated China 1–0 to advance to the semifinals of the 2015 FIFA Women’s World Cup. Likewise, the lexical nominalization half-pounder, based on the noun half-pound and is used in an expression like Give me a half-pounder, may denote a hamburger in a fast-food restaurant, a can of tobacco in a smoke shop, a bag of jelly beans in a candy shop, or a steelhead trout among fishing aficionados.

While many lexical nominalizations tend to have more uniform denotations, grammatical nominalizations do not have fixed or uniform denotations, and speech context plays an important role in determining and selecting the denotation/reference most consistent with the context. For example, the Spanish grammatical nominalization [el [que Ø es blanco]_{nmlz}] (the [nmlzr Ø is white]) ‘the one which is white’ can refer to a range of objects classed as masculine matching its denotation of an entity that is white. In actual usage, the context and the Gricean Cooperative Principle determine the reference. So, El que es blanco would be understood to be referring to a white car when uttered in response to the question ¿Qué coche te gusta? “Which car do you like?” and a white hat when it answers the question ¿Cuál sombrero usarás hoy? “Which hat will you wear today?”. Our claim is that there is nothing like a deletion of a head noun or a pronominal element involved here. The construction [el [que Ø es blanco]] is a complete structure, a grammatical argument nominalization (see Section 3.2), whose reference in discourse is determined by the context, exactly like the determination of the actual reference of a metonymic expression such as the United States following questions like “Who has decided to pull out from the Paris agreement?” or “Who defeated China 1-0 to advance to the semifinals of the 2015 FIFA Women’s World Cup?” or the like.

For the purposes of this paper, we offer the following definition of nominalization:
Nominalization is a metonymy-based grammatical derivation process yielding constructions associated with a denotation comprised of entity (thing-like) concepts that are metonymically evoked by the nominalization structures, such as events, facts, propositions, resultant products and event participants. Nominalizations, as grammatical structures, are similar to nouns by virtue of their association with an entity-concept denotation; they both denote thing-like concepts, which provide a basis for the referential function of an NP headed by these nominals.

Notice at this juncture that nominalization is not a morphosyntactic notion. In particular, nominalization structures may or may not have associated morphology, as the comparison between the two relevant forms below indicates.

\[
\begin{align*}
(2-1) & \quad a. \ drive & > & \text{driver} \\
& \quad b. \ cook & > & (a) \ cook
\end{align*}
\]

\[
\begin{align*}
(2-2) & \quad a. \text{Quechua (Cuzco dialect; Lefebvre and Muysken 1988)} \\
& \quad \text{Juan potato-acc eat-NMLR-3-ACC know-1sg} \\
& \quad \text{[Xwan papa-ta mikh-sqa-n]_{NMLZ}^{NP} ta yacha-ni. (lit.) 'I know Juan('s) eating potatoes.'} \\
& \quad b. \text{I know } ([Juan eats potatoes]_{NMLZ}^{NMLR})_{NP}
\end{align*}
\]

The noun \textit{cook} has no marking indicating that it is a nominalized from like the noun \textit{driver} with a nominalization marker, yet \textit{cook} and \textit{driver} are functionally alike both semantically and syntactically. It is on the basis of this functional similarity that we treat the nouns \textit{driver} and \textit{cook} as nominalized forms despite the difference in morphology. The same logic applies to the relevant structures in (2-2a) and (2-2b); semantically both denote a fact, and syntactically both function as an object argument of the verb meaning “know”. Treating the relevant structure in (2-2a) as a nominalization, while not recognizing a nominalization in (2-2b) on account of the presence and absence of nominalization morphology is like treating the Quechua form \textit{wanu-či} (die-CAUS) ‘kill’ as causative but not the English verb \textit{kill}. Just as causation is not a morphosyntactic notion, nominalization is not a morphosyntactically definable phenomenon.

2.2 Meaning range and the nature of lexical nominalizations

As seen above, metonymy allows various form-meaning connections beyond the conventional lexicon of basic, underived nouns. Nominalization as a grammatical process yields forms (nouns and larger structures or constructions), which are called “nominalizations” and are labeled as \[\ldots\]_{NMLZ} in this paper, that denote

---

5. Labelling the derivation \textit{cook} > (a) \textit{cook} as “conversion” is besides the point.
things and thing-like entities evoked by the derived nominalization structures. Verbal-based nominalizations, as noted above, evoke concepts intimately related to what the verbal bases denote, namely states (slowness, freezing), processes (flowing, freezing), activities (fight, skating, building, employment), facts, propositions, attendant protagonists such as agents (the noun cook, employer) and patients (fryer, keeper “a fish that is of sufficient size to be caught and retained without violating the law”, employee), resultant products (the nouns freeze, building, painting, writing), as well as instruments (plier, screwdriver) and locations (landing, bus stop) inherently or conventionally associated with particular events. While some lexical nominalizations involve morphology that delimits the range of meanings associated with the derived nominals, as in the case of the English suffix -er and others seen above, some others may form nominalizations with a greater range of meanings as with the so-called gerundive -ing form in English. Of the Japanese lexical nominalizations, stem nominalizations, involving -i/-Ø suffix, display a diverse array of meaning patterns on their own, but more typically in forming compounds with another nominal element. However, as in the English cases discussed above, the form-meaning connections are not random and are metonymically bound such that some meaning patterns are more consistently observed while others are not. What follows summarizes major form-meaning patterns that Japanese stem nominalizations display.


e. Patient: yatoi ‘employee’, tukai ‘errand runner’, tumami ‘what is picked/hors d’oeuvre’, ture ‘one taken along/companion’

6. These nominalizations, often labeled ‘infinitive’, involve the suffix -i attaching to a consonant-final root/stem (e.g. odor- > odor-i ‘dancing/dance’), and -Ø to a vowel-final root/stem (e.g. kake- > kake ‘betting’). See Shibatani (2018b) for discussions of these lexical nominalizations and their theoretical implications, including critiques of earlier treatments of them.
What we observe in the remainder of this paper is that grammatical nominalizations to a large extent parallel these lexical nominalizations in the range of meaning extension, indicating that these two are closely connected phenomena, a fact that is also indicated by morphology in many languages (see Section 2.4). Indeed, there are historical connections between grammatical nominalizations and lexical nominalizations such that the latter often arise from the former (see Fleck, this volume, for an extensive discussion on this). The fact that the Japanese stem nominalization seen above displays such a diverse range of meaning, rather than more narrowly circumscribed patterns as in the case of employment, employer and employee, indicates that it was once a productive grammatical nominalization process (see Shibatani (2018a) for the synchronic data suggesting this and Section 5.1 below, where so-called clause-chains are discussed).

Despite these connections between lexical and grammatical nominalizations, there can be differences between the two. While the above exposition takes a formal difference as a criterion for distinguishing between lexical (those being single words, possibly with internal structural complexity as in the case of compound words and those containing vestiges of certain affixes) and grammatical nominalizations (those having phrasal structures larger in size than words), the distinction between these two types of nominalization can be quite subtle and difficult to pinpoint in many cases. Since this paper is mainly concerned with grammatical nominalizations, we spend the rest of this subsection discussing the
nature and issues pertaining to what may be called lexical nominalizations and the
distinctions between them and grammatical nominalizations.

One difference observed across languages is a formal/morphological differ-
ence. Lexical nominalizations are often associated with specific morphology, as in
*employ > employment*, but grammatical nominalizations may, in many languages,
show no morphological marker at all, as in *I know [John employs Bill]*[^nmzl] where
the nominalization denoting a fact has the same *internal structure* as the sentence
*John employs Bill*. A formal characteristic of grammatical nominalizations like this
has led many to view them as clauses or even sentences. But, as noted above, even
lexical nominalizations may not have any morphological indication, e.g. *cook > a
cook, walk > a walk*, and thus the notion of nominalization is independent from
morphological marking.

Another characteristic that distinguishes lexical nominalizations from gram-
matical nominalizations is that the former have irregular gaps in the meaning pat-
terns. Returning to the Japanese stem nominalizations discussed above, many of
them based on action verb roots allow both activity and agent/instrument read-
ings (*suri* ‘pickpocketing/pickpocket’, *hitō-gorō* ‘manslaughter/killer’, *simi-nuki*
‘stain removing/stain remover’), whereas many other similar forms have only one
reading. Forms like *yama-nobori* ‘mountain climbing’, *uo-turi* ‘fish catching’, and
*sumi-yaki* ‘charcoal-making’ only denote activities, whereas *uta-utai* (song-sing-
ing) ‘singer’, *e-kaki* (picture-drawing) ‘painter’, and *sumō-tori* (sumo-taking) ‘su-
mo-wrestler’ name only agents and not activities such that while *[yama-nobori]-
suru* ‘do mountain-climbing’ is possible, *[uta-utai]-suru* ‘do song-singing’ is not.

Finally, grammatical nominalizations differ from lexical nominalizations
in that their meanings tend to be compositional, while lexical nominalizations
having a word status may undergo meaning specialization, as in the case of *e-
kaki* (picture-drawing) ‘a painter’ denoting a professional artist, as opposed to the
grammatical counterpart *asoko de e o kaku no* (there loc picture acc draw npm)
‘(one) who draws a picture there’, which is free from such a conventionalized re-
strictive meaning. The term “lexicalization” is sometimes used to refer to this kind
of specialization in meaning. In the case of *e-kaki* ‘a painter’ above, this sense of
lexicalization coincides with the status of the form as a noun. But since there are
phrases whose meanings are specialized/noncompositional, namely idioms (e.g.
*all hell breaks loose, kick the bucket*), meaning specialization itself does not define
words. And there are many lexical nominalizations that do not have specialized

[^nmzl]: Bear in mind that structural resemblances do not guarantee that we are dealing with the same
grammatical units. In particular, we reject a view that certain nominalizations are clauses/sen-
tences because their internal structural properties are similar to those of clauses and sentences:
(see Section 6.2 on the definitions of clauses, sentences, and nominalizations).
meanings. Indeed, there are cases where a nominalized form may convey both lexicalized meaning and literal, compositional meanings. Consider the following forms from Central Alaskan Yup’ik:

(2-4) Central Alaskan Yup’ik (Tamura 2017 quoting Jacobson 1984: 450, 560)

i. *pi- ‘do’ > *pi-sta ‘doer, servant’
   *pi-sta-i ‘the one who did something to him, ‘his servant’

ii. *angu- ‘pull’ > *angu-n ‘the one who is pulling, ‘man’

iii. *mikel- ‘small’ > *mikel-nguk ‘the one who is small, ‘child’

iv. *kipus- ‘buy’ > *kipus-vik ‘the place for buying, ‘store’

Since meaning does not provide a definitive criterion for lexical categories (cf. resemble vs. similar in English, where the former is a verb and the latter an adjective), the morphosyntactic status of the form in question must be ascertained in order to determine its lexical (word-unit) status. If the form has morphosyntactic properties of ordinary nouns, then it is a lexical nominalization. In the case of singing in English, for example, we must recognize two distinct forms. One of them occurs as a word by itself, as in her beautiful singing (impressed us). The other singing does not, e.g. *her beautifully singing (was quite a show). This singing occurs only in a phrasal form, as in her beautifully singing the national anthem (was quite a show). Compare this with the former singing, which is modified by an adjective and which does not form a phrase with a noun phrase, as in her beautiful singing *(of) the national anthem, where the preposition of is required. Singing the national anthem is not a word and therefore it is a grammatical nominalization, while singing as in singing of the national anthem is a word and displays many of the essential properties of ordinary nouns. However, the grammatical nominalization singing the national anthem also displays some major properties of ordinary nouns, such as being modified by a genitive determiner, as in her singing the national anthem, and it heads a subject and an object NP, as in [singing the national anthem] would be quite appropriate, yet it does not allow the marking by an article (*a/*the singing the national anthem), unlike ordinary nouns or the lexical counterpart (a/the singing of the national anthem).

As the above discussion demonstrates, the question of whether a nominalization structure is a noun (lexical) or not boils down to the degree of similarity of the form to ordinary nouns. It was on observations like this that John Robert Ross (aka Haj Ross) proposed what he called the nouniness squish, capturing the cline of nouniness among various types of nominals:

---

9. There are also cases where English does not have corresponding nouns and descriptive grammars must resort to an analytic translation such as ‘one who does X’.
that clauses > for to clauses > embedded questions > Acc ing complements > poss ing complements > action nominals\textsuperscript{10} > derived nominals > underived nominals (Ross 1973)

One may use the term “lexical nominalization” in reference to those forms that display the total or a majority of morphosyntactic properties of ordinary nouns and “grammatical nominalization” for those divergent from ordinary nouns yet showing varying degrees of nouniness in morphosyntax. These two types of nominalizations are best considered to form a continuum on formal grounds. What unifies them (and their subtypes) as nominalizations is their semantic function (denoting things and thing-like entities), with varying degrees of morphosyntactic repercussions of this function depending on their types.

In view of the infeasibility of applying nouniness tests to the data from a diverse array of languages below, the tactic adopted in this paper is to treat those forms discussed under “nominalization” in the literature as lexical nominalizations, while treating as grammatical nominalizations those that are discussed elsewhere in the grammar under such headings as “relative clauses” or “subordination”, which tend to include verbal morphosyntactic properties (e.g. person and voice marking, the presence of an argument NP and an adverbial modifier) not associated with ordinary nouns (and “true” lexical nominalizations). It is interesting and relevant to examine if the two readings for each of the Yup’ik forms in (2–4) above correlate with differences in morphosyntactic properties.\textsuperscript{11} Either way, we would not know whether a given form is a lexical or grammatical nominalization until we apply available morphosyntactic nouniness tests and determine the degree to which the given form resembles ordinary nouns in the language.\textsuperscript{12}

\textsuperscript{10} The singing of singing of the national anthem appears to retain a certain degree of “verbiness” in that, while Her beautifully singing of the national anthem impressed them is totally ungrammatical, Her singing of the national anthem beautifully impressed them does not appear to be totally ungrammatical.

\textsuperscript{11} Fleck’s contribution to this volume exactly addresses this important issue and shows that the formally identical forms display a different external property (whether or not the nominalized forms can be modified by a possessive form) depending on whether they are associated with a conventionalized meaning denoting a specific object (e.g. a stick or club) or with a schematic meaning denoting a variety of objects (e.g. any object that one may use to hit something/someone) (see also Tamura, this volume).

\textsuperscript{12} More can be said about this distinction from various theoretical perspectives such as the distinction between lexicon/morphology and syntax, as in the Generative Grammar framework. The framework that better accommodates the lexical-grammatical continuum is Langacker’s Cognitive Grammar (Langacker 1987, 1991) that recognizes degrees of entrenchment and schematicity, both of which attempt to capture the degree of productivity of construction types and the associated difference in meaning (idiomaticity/compositional). We are unable to go into
2.3 Inputs to nominalization processes

Perhaps influenced by the term “nominalization”, which suggests turning something in one state into something in another state, many researchers, as indicated by the quotes above, believe that nominalization is restricted to derivations that turn verbal inputs into nominal outputs, as in Payne’s (1997) and Malchukov’s (2004) definitions or to those that take members of non-nominal categories as their inputs, as in Crystal’s (1980) definitions. While recognizing nominalizations based on nominal inputs, Comrie and Thompson (1997/2007) give short shrift to such cases by allocating only one page and a few additional lines in their 47-page discussions on lexical nominalizations. A similar bias toward verbal-based nominalization is also clearly seen in all the papers collected in Yap, Grunow-Hårsta & Wrona (2011).

It is unclear why these researchers have decided to focus more on verbal-based nominalizations, when even such a well-known language as English exhibits nominal-based nominalizations. The case in point involves the so-called agentive suffix -er, deriving verbal-based forms such as play > player, sing > singer, which, everyone would agree, is a clear case of lexical nominalization. But this process takes a wide range of nominal inputs, as demonstrated by villager, New Yorker, rear-ender, right-winger, leftfielder, knuckleballer, tenner, 49ers, lifer, spring breaker, the aforementioned half-pounder, 18-wheeler, etc. While many of these are not strictly agentive, they denote entities that are closely associated with the meaning of the base forms. Whether a derived form denotes an agent or non-agentive entity simply depends on the nature of the base form; verb-based nominalizations denote an entity most closely associated with activities, namely an agent (and possibly an instrument and others), whereas nominal-based ones denote other types of entities metonymically evoked in close association with the denotations of the base nouns, such as the people associated with specific locations one way or another, and those entities associated with a specific quantity, time, or manner.

The above pattern certainly is not limited to English. Parkatêjê, a Je language in northern Brazil, has the agentive suffix -katê, which nominalizes verb roots (e.g. krere ‘sing’ > krere-katê ‘singer’; jakre ‘write’ > jakre-katê ‘writer’). But this suffix productively applies to animal names as well, producing forms like pryre ‘animal’ > pryre-katê ‘(animal) hunter’, rop ‘jaguar’ > rop-katê ‘jaguar hunter’, and kukryt ‘tapir’ > kukryt-katê ‘tapir hunter’ (author’s field notes).

these alternative perspectives due to other empirical, and more pressing, issues that need to be addressed, but how the difference such as the degree of productivity and meaning compositionality correlates with the degree of formal nouniness discussed here would be an interesting and perhaps worthwhile topic to pursue.
Central Alaskan Yup’ik has several nominalizers that apply to verbs to derive nouns (lexical nominalizations) and grammatical nominalizations. But most, if not all, also take nouns as inputs.

(2-5) Locative/Time nominalizer -vik ‘place to V; place or time of V-ing; place for N’ (Tamura 2017 quoting Jacobson 1984: 586; Tamura this volume)
   a. Verb-based lexical nominalization
      i. nere- ‘eat’ > nervik ‘dining hall, restaurant, table’
      ii. qavar- ‘sleep’ > qavarvik ‘sleeping bag, bedroom’
   b. Noun-based lexical nominalization
      qimugta ‘dog’ > qimugtevik ‘kennel, dog house’

(2-6) Instrumental nominalizer -cuun ‘device for V-ing, device associated with N’ (Tamura 2017 quoting Jacobson 1984: 450)
   a. Verb-based lexical nominalization
      i. mingqe- ‘saw’ > mingqescuun ‘sawing machine’
      ii. igar- ‘write’ > igarcuun ‘pencil, pen’
      iii. nere- ‘eat’ > nerrscuun ‘fork, eating utensil’
   b. Noun-based lexical nominalization
      i. anuqa ‘storm’ > anuqessuun ‘wind generator, storm lantern’
      ii. arnassagaq ‘old woman’ > arnassagarcuun ‘old-age pension for a woman’

Yagua in northwestern Amazonia has nominalizing classifiers that apply not only to verbal roots but also to adjectival as well as nominal roots (e.g. ɨtyọ–jay (sleep-clf.pelt) ‘sleeping mat’, ɨqa uomo–daisit (big-clf.thin.pole) ‘big blowgun, pole’, ɨqọsq–jaq (light-clf.liquid) ‘kerosene’) (Payne 1985). The Salish language Halkomelem has similar nominalizing classifiers that also apply to verbal, adjectival, and nominal roots (e.g. ɨtət=ɨwtxw (sleep=clf.house) ‘hotel, bedroom’, qaq’iy =e’wtxw (sick=clf.house) ‘hospital’, tel=e’wtxw (money=clf.house) ‘bank’) (Gerds and Hinkson 2004).

Languages outside of the Americas also allow nominalization to apply to nouns. Gá, a Kwa language spoken in Ghana, has the “agentive” suffix -lɔ, which nominalizes verbs, as in jù ‘steal’ > jù-lɔ ‘thief’ and tsɔ–sɔ ‘show; teach’ > tsɔ–sɔ–lɔ ‘teacher’. Just like the English and the Parkatêjê counterpart, this suffix also applies to nouns, as in ənɪhɔ ‘laziness’ > ənɪhɔ–lɔ ‘lazy person’, bèli ‘quarrel(n.)’ > bèli–lɔ ‘quarrelsome person’, and əməlɛ ‘lie (n.)’ > əməlɛ–lɔ ‘liar’. Gá also has the “agentive” nominalizer -tʃɛ, which applies to both adjectives and nouns; əgbɔ ‘big’ > əgbɔ–tʃɛ
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization


Chinese has a number of agentive suffixes that derive nouns from verbs, which are similar in function to the English -er suffix seen above. They are also similar in that they apply to noun inputs as well, as shown in the Mandarin forms below:

(2-7) V > N N > N
-zhě (者 ‘person’); jì-zhě
to record-suf ‘reporter’
zuò-zhě
to make-suf ‘author’
zhàng-zhě
to be elderly-suf ‘elderly man’
-şhōu (手 ‘hand’); zhù-şhōu
to help-suf ‘assistant’
hào-şhōu
to be good-suf ‘skilled person’
xióng-şhōu
to be ferocious-suf ‘murderer’
-jiā (家 ‘house’); zuò-jiā ‘writer’
to make-suf ‘writer’
huà-jiā
to paint-suf ‘painter’
pīpīng-jiā
to criticize-suf ‘critic’

In Section 7 we offer a novel analysis of the genitive/possessive form (e.g. my, mine) as an instance of nominal-based nominalizations. As a way of summarizing the discussions above, observe the following examples from the Athabaskan language
Slave and the Austronesian language Tagalog that shed light on two important issues that are pursued in this paper. Namely, (i) that nominalization applies to nominals as well (the point proven above), and (ii) that lexical nominalizations may share morphology with grammatical nominalizations indicating that the two constitute a unified phenomenon. The Slave nominalizing suffix -i and its phonological variants productively derive verbal-based grammatical nominalizations that denote an event protagonist (e.g. (2-8c)). It also derives lexical nominalizations (e.g. (2-8a)), and it applies to numerals as well, which are assumed to be nouns in this language, as in most other languages (e.g. (2-8b)).

(2-8)  
Salve (Rice 1989)  
\begin{enumerate}
    \item Verbal-based lexical nominalization
        \begin{enumerate}
            \item \textit{\textbullet}ehdzo.i
                \begin{itemize}
                    \item trap.nmlzr
                    \end{itemize}
                \begin{itemize}
                    \item ‘a trap’
                \end{itemize}
            \end{enumerate}
            \begin{itemize}
                \item cf. \textit{\textbullet}ehdzo ‘S/he traps something.’
            \end{itemize}
    \item Nominal-based nominalization
        \begin{enumerate}
            \item \textit{dį-i} \textit{\textbullet}ehk’ē
                \begin{itemize}
                    \item four-nmlzr 3.shot
                \end{itemize}
            \begin{itemize}
                \item ‘S/he shot four \{ANIMALs\}.’
            \end{itemize}
        \end{enumerate}
    \item Nominal-based nominalization
        \begin{enumerate}
            \item \textit{lake’e–e} \textit{ragokedēhwe}
                \begin{itemize}
                    \item five-nmlzr 3.start back
                \end{itemize}
            \begin{itemize}
                \item ‘Five \{PEOPLE\} started back.’
            \end{itemize}
        \end{enumerate}
\end{enumerate}
\begin{enumerate}
    \item Verbal-based grammatical nominalization
        \begin{enumerate}
            \item \textit{[nįwā kedaw’i] i \textbullet}goghāeyida
                \begin{itemize}
                    \item long 3pl.sat nmlzr pl 1sg.saw.3pl
                \end{itemize}
            \begin{itemize}
                \item ‘I met ones who stayed a long time.’
            \end{itemize}
        \end{enumerate}
\end{enumerate}

In a similar vein, the Tagalog locative nominalizer -(\textbullet)an, which productively derives verbal-based locative grammatical nominalizations, also attaches to nouns and yields new nouns (lexical nominalizations) denoting locations that the referents of the base nouns are conventionally associated with (Schachter & Otanes 1972/1983: 98ff).

\textsuperscript{13} The glosses in the examples from other sources are mostly original, except for what I consider to be nominalizing morphemes, which in the originals are glossed variously as NOM, N, REL, COMP, SBR (subordinator) etc. I take the liberty of glossing them uniformly as nmlzr. The nominalization structures, on the other hand, are marked as […]nmlz. It is amusing indeed to notice that Comrie (2006) relabels as rel(!) the Tibetan nominalizers so recognized and glossed as NOM by the Tibetan specialist Scott DeLancey (DeLancey 2002).
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(2-9) Tagalog nominal-based lexical nominalization
i. aklat ‘book’ > aklat-an ‘library’
ii. halaman ‘plant’ > halaman-an ‘garden’
iii. tarangka ‘lock’ > tarangka-han ‘gate’

(2-10) Tagalog verbal-based grammatical nominalization
Ang [b < in > ili-han ng lalaki ng saging] ang tindahan ni Dodong.
Top buy<Pfv>-LF GEN man GEN banana TOP store GEN Dodong
‘Dodong’s store is where the man bought the bananas.’

Having demonstrated that nominalization applies to nominals and having taken a quick view on the possibility that nominalization morphology produces units larger than words, we shall now more closely look at morphological connections between lexical and grammatical nominalizations as a way to motivate the recognition of grammatical nominalizations as such.

2.4 Outputs of nominalization processes

While Payne’s (1997) and Comrie and Thompson’s (1985/2007) definition of nominalization restricts the outputs to nouns, and the translations of the term “nominalization” in some grammatical traditions may literally mean noun-formation, as the term meishi-ka in Japanese linguistics does, the process actually creates units larger in size than words, as we have already seen above. The Slave form [[nįwą́ kedaw’į] i] ‘one who stayed long’ and the Tagalog form ang [b < in>ili-han ng=lalaki ng=saging] ‘where the man bought bananas’ are cases in point. Our arguments for treating these as (grammatical) nominalizations, rather than, say, verbal phrases, rest on two crucial facts pertaining to these structures. The foremost is the fact that these phrasal structures denote substantive concepts just like ordinary nouns and lexical nominalizations. We shall dwell on this fundamental property of grammatical nominalizations throughout this paper, but especially in Section 6.2, where we discuss differences among nominalizations, clauses, and sentences.

The other fact that motivates analyzing the structures like the Slave and Tagalog grammatical nominalizations above as such is that they share morphological marking with uncontroversial lexical nominalizations. This kind of morphological connections between lexical and grammatical nominalizations are not limited to Slave and Tagalog, as we saw earlier in terms of the forms her singing of the national anthem and her singing the national anthem, although many other languages show different marking patterns for the two types of nominalization. While nominalization is not definable in terms of morphology, morphological connections between lexical nominalizations and grammatical nominalizations provide concrete and positive evidence that the latter represent a nominalization
phenomenon. Below we offer a sampling of languages from a diverse array of language families in the Americas and a few additional examples from outside this region to show that our treatment of grammatical nominalizations as such is also morphologically supported.

Let us start with languages in South America beginning with the Tapiete examples below.

(2-11) Tapiete
(Tupí-Guaraní; Bolivia, Paraguay, Argentina; González 2005; Ciccone 2008)

a. Lexical nominalizations
   i. hē’ē ‘be sweet’ > hē’ē-wā ‘sugar, honey’
   ii. arika’ē ‘long ago’ > arikā’ē-wa ‘ancestors, history’

b. Grammatical nominalizations
   i. [mbiri-iwi ou-wa] kō’ēn-wai ou far–dir 3.come-nmlzr in the morning come
   ‘(The one) who comes from far away has arrived this morning.’
   ii. [kā’ā tenta-pe hau-wa] hayasi yesterday town-loc 1.eat-nmlzr be.rotten
   ‘What I ate yesterday in town was rotten.’

The first grammatical nominalization above is a subject argument nominalization that denotes an agentive entity (similar to the English agentive lexical nominalization employer) metonymically evoked by the nominalization structure marked by the nominalizer -wa, which also derives lexical nominalizations. The second form, on the other hand, is an object grammatical nominalization denoting a patientive entity (similar to the English patientive lexical nominalization employee).

Kakataibo has several different nominalizers for lexical and grammatical nominalization but the two processes do share the nominalizer -kë, similarly to the Tapiete nominalizer illustrated above. The first grammatical nominalization is a patient argument nominalization like the patient lexical nominalization tua-kë ‘son of a woman’, while the second is an event nominalization that metonymically denotes a fact pertaining to the event.

(2-12) Kakataibo (Panoan; Peru; Zariquiey 2011: 297, 632, 638)

a. Lexical nominalization
   i. mapun- ‘to cover’ > mapun-kë ‘house’
   ii. tua- ‘to give birth’ > tua-kë ‘son of a woman’

b. Grammatical nominalization
   i. ashi ka ēn ñui-kas-kë ‘iashin
   a=ishi ka [‘ē=n ñui-kas-kë] ‘i-a-x-ín
   that=only NAR.3P 1sg=A tell-DES-nmlzr be-perf-3p-prox
   ‘Only that was what I wanted to tell.’
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(2-13) Shawi (Kawapanan)

a. Lexical nominalization
   i. shawi-napi ‘the gossiper’
   ii. ni’i-napi ‘the watcher’
   iii. u’u-napi ‘drinker’

b. Grammatical nominalization
   atari nusha a-ka’n-napi
   chicken meat CAUS-eat-NMLZR
   ‘the one who makes someone eat chicken’

The next examples come from Bolivian Quechua.

(2-14) Bolivian Quechua (Bills et al. 1971: 101, 106)

a. Lexical nominalization
   i. llank’a-q ‘worker’
   ii. yanapa-q ‘helper’
   iii. ranti-q-kuna ‘buyers’
   iv. wayk’u-q ‘cook’

b. Grammatical nominalization
   i. haway runa [qulqi muna-q] (based on Bill’s et al. 1971: 274)
      that  man money want-NMLZR
      ‘That man is the one who wants the money.’
   ii. Munaku-ni [runa maqa-q]
      love-1SG man hit-NMLZR
      ‘I love the one who hit the man.’

Turning to North America, we again observe widespread morphological connections between lexical and grammatical nominalizations demanding recognition of the latter as a type of nominalization.

(2-15) Northern Paiute (Numic, Uto-Aztecan; Thornes 2003: 118, 431)

a. Lexical nominalization
   i. tiničuí-di
      teach-NMLZR
      ‘teacher’

Rojas-Berscia’s contribution to this volume also shows that the Kawapanan language Shawi derives both lexical and grammatical nominalizations by the use of the nominalizer -napi, as seen below.

‘(The fact) that Juan works in the hotel is good.’
ii. hoawai-di
   hunt-NMLZR
   ‘hunter’

iii. nayakwi-di
   (play)handgame-NMLZR
   ‘gambler’

b. Grammatical nominalization

i. oʔo iwa-ʔyu su=umu-ʔyu naʔa
dem many-NOM NOM=that.kind-NOM grow
su=[na-tihona-di]
NOM=MM-dig.roots-NMLZR
‘There’s a lot of that kind growing out there for the digging.’ (i.e. that
which is dug).

ii. ta=sakwa umi-no [na-kwii-di]
1DL.INCL=MOD 3PL- COM MM-smoke-NMLZR
‘We should (go) with them to the smoke-pit.’ (i.e. that which
smokes)

Shoshone, which also belongs to the Numic branch of Uto-Aztecan family, like
Northern Paiute, has several types of nominalizers used for lexical and grammati-

Kiowa, forming the Kiowa-Tanoan family with the Tanoan pueblos of New
Mexico and Arizona, has a nominalizer that derives nouns from indefinite and
interrogative roots as well as from some locative roots and phrases, according to
Watkins (1984). This suffix, as seen below, appears to have a much more produc-
tive use as a nominalizer deriving grammatical nominalizations, which are used as
modifiers of nouns-in so-called relative clause constructions-as in other languages.

(2-16) Kiowa (Kiowa-Tanoan; Watkins 1984: 108, 109, 230)

a. Lexical nominalization

i. hôn-dé ‘someone’, ‘something’ ‘some kind of’ (< hôn- indefinite root)
ii. hôn-dé ‘who?, what? ’ ‘what kind of?’ (< hôn interrogative root)
iii. tô-yô-dé ‘lid, cover’ (< tô-y ‘atop’)
iv. tôy-dé ‘household goods, furniture’ (< tô-y ‘in the house’)
v. giorni-té ‘other(s), different one(s), outsider(s)’ (gû-y ‘outside’)

b. Grammatical nominalization

i. óy-gô k’yáhí’i [ô-dé-dé] k’yátta’y-k’i’ô-dô
there-ADV man 3SG-stand-NMLZR chief-male 3SG-be
‘The man (who is) standing there is a chief.’
ii. \( k'\dot{r} \quad [k'\dot{d}a\dot{l}-\dot{y} \quad o-\dot{d}l + s\dot{s}l-d\dot{e}] \)
\( \) wood \( \) wagon-on \( 3sg-load+be-MMLZR \)
\( gy\dot{a}-p'\dot{e}\text{tt} \)
\( 1sg/agt:sgobj-take=down-impf \)

'I am unloading wood that was loaded in the wagon.'

The Siouan language Crow spoken in Montana has, among others, the agentive (\( ak- \)) and the locative/temporal/manner (\( ala- \)) nominalizers that derive lexical nominalizations. These are also used in grammatical nominalizations in a parallel manner, as below.

(2-17) Crow (Siouan; Graczyk 2007: 254, 255)

a. Lexical nominalization
   i. \( ak\text{-}diss\dot{h}i \) ‘dancer’ < \( diss\dot{h}i \) ‘dance’
   ii. \( ak\text{-}kum\dot{m}i \) ‘singer’ < \( kum\dot{m}i \) ‘sing’
   iii. \( ala\text{-}chiwakåå\text{-}u \) ‘church’ < \( chiwakìì + pl \) (‘where they pray’)
   iv. \( ala\text{-}såhta \) ‘fork in a river’ < \( såhta \) ‘forked, pronged’

b. Grammatical nominalization
   i. \( hileen \quad [ak\text{-}iss\dot{h}i\text{-}ss\text{-}aa\text{-}lee\text{-}sh] \quad awe\text{-}taa \quad xëmm\text{-}ak \)
      these \( \) NMLZR-top-goal-port-go-det ground-path lie-ss
      dupesaa-(a)-ahk-uu-k
      ‘The ones who had brought him to the top were lying on the ground panting’
   ii. \( pü\dot{a}ee \quad [bale \quad ala\text{-}satchë] \quad ko \quad kukaå \quad húu\text{-}ssaa\text{-}k \)
      smoke \( \) wood NMLZR-thick PRO source come-NEG-DECL
      ‘The smoke isn’t coming where the trees are thick [the forest].’

A close parallel to the Crow pattern is seen in the Yuman language Jamul Tiipay in southern California, in which lexical nominalizations and grammatical nominalizations share nominalization elements (prefixes, suffixes, length ablaut), as observed below.

(2-18) Jamul Tiipay (Yuman; Miller 2001: 117, 118, 122, 214)

a. Lexical nominalization
   i. \( kwe\text{-}ch\text{-}cheyaaw \) ‘singer’ < \( cheyaw \) ‘to sing’ (Subject nominalization)
   ii. \( ke\text{-}meyaally \) ‘tortilla maker’ < \( meyally \) ‘to make tortillas’ (Subject nominalization)
   iii. \( ya\text{'-}ma\text{-}ch \) ‘powwow’ < \( iima \) ‘to dance’ (Oblique nominalization)
   iv. \( a\text{-}sii\text{-}ch \) ‘liquor’ < \( si \) ‘to drink’ (Oblique nominalization)
b. Grammatical nominalization
   i. `[met’aiar-i ke-pa-ch]-pu yaach Uuyaaw` outdoors-loc nmlzr-be.prs.srs-nmlzr-dem 1.sj know
      (Subject nominalization)
      ‘I know the one who is out there.’
   ii. `[nya’ru me’a-y stu-ch]-pu nyaach uuyaaw` money where-loc pick.up-nmlzr 1.sub know
       (Oblique nominalization)
       ‘I know where you got that money=I know the place where you got the money.’

Salish languages of the Pacific north coast all appear to contain the morpheme `s-` that produces both lexical and grammatical nominalizations, as in Musqueam below.

(2-19) Musqueam (Halkomelem, Salish; Suttles 2004: 77, 101 264, 265)
   a. Lexical nominalization
      i. `s-ziłtn‘food’ < ziłtn‘eat’
      ii. `s-ziłθom‘clothing’ < ziłθom‘get dressed’
      iii. `s-pìw‘ice’ < pìwet‘freeze it’
      iv. `s-kʷix‘name’ < kʷixət‘name it’
      v. `s-yá’ys‘work’ < yá’ys‘work’
   b. Grammatical nominalizations
      i. `kʷθɔ [niʔ n̓ə-s-ziłxi-t]`
         art aux my-nmlzr-borrow-tr
         ‘what I rent him’
      ii. `tòy [k’wɔ [s̚-q’en-əθst-ct]]`
         good art nmlzr-return-self-our
         ‘We’d better return.’ (lit. Our returning would be good.)

Finally, Eskimo languages display several nominalizers that yield different types of nominalizations such as agentive, instrumental, and locative nominalizations. The following illustrates the use of the agentive nominalizer -(s)ta in Central Alaskan Yup’ik marking both lexical and grammatical nominalizations.

(2-20) Central Alaskan Yup’ik (Miyaoka 2012: 532, 533, 535)
   a. Lexical nominalization
      i. `cali-sta‘worker’
      ii. `kuvya-sta‘one who fishes by net’
      iii. `ikayur-ta‘helper’
   b. Grammatical nominalization
      i. `[neqe-m nere-sti-i]`
         fish-rel.sg eat-nmlzr-abs
         ‘one that eats the fish’ cf. `qimugta ner’-uq neq-mek`
         ‘the dog is eating fish’
ii. carayi-i-m     tangvag-ti-i  
ghost-EV-REL.SG see-NMLZR-ABS.3SG.SG  
‘the one who saw the ghost’

Just to complete the picture, let us look at a few languages outside the Americas. Thai (Tai-Kadai) has the nominalizer thii, which marks lexical nominalizations, many of which form noun compounds, as well as grammatical nominalizations of both those that denote event protagonists and those that denote state of affairs or facts.

(2-21) Thai (Tai-Kadai)
a. Lexical nominalization
i. thii-nâŋ  
NMLZR-sit  
‘seat’
ii. thii-tât    lép  
NMLZR-cut nail  
(Iwasaki & Ingkaphirom 2005: 45)  
‘nail clipper’
iii. thii-lâaw    dinsɔ̌ɔ  
NMLZR-sharpen pencil  
‘pencil sharpener’
iv. thii-pɔt    krapɔŋ  
NMLZR-open can  
‘can opener’
b. Grammatical nominalization
i. chǎn cà?    sàây   [thii    khwɛɛn nay tûu]  
I will wear NMLZR hang in closet  
(courtesy of Kingkarn Thepkanjana)  
‘I will wear the one that hangs in the closet.’
ii. tɛ̀-wâa dîi    nà   [thii    mây mii    khay pen alay]NMLZ  
but good PP NMLZR NEG have who CPP what  
(Iwasaki & Ingkaphirom 2005: 255)  
‘But it was good that no one was hurt.’

Korean has a fair number of nouns derived from verbs by -um suffixation. This suffix is also involved in grammatical nominalizations, which are somewhat archaic sounding, according to the Korean speakers consulted.
(2-22) Korean (Courtesy of Sung-Yeo Chung)
a. Lexical nominalization
   i. cwuk- ‘die’ > cwuk-um ‘death’, cwuk-em ‘corpse’
   ii. mwut- ‘bury’ > mwut-um ‘burial’, mwut-em ‘grave’
   iii. kuli- ‘draw’ > kuli-m ‘drawing, picture’
b. Grammatical nominalization
   [mayil swul-ul masi-m]-un [kenkang-ul ilh-um]-ul
   every.day alcohol-acc drink-NMLZR-TOP health-acc lose-NMLZR-ACC
   [uymiha-m]-i-yo.
   mean-NMLZR-COP-ASSERTIVE
   ‘To drink alcohol every day means to lose one’s health.’

As final examples for morphological connections between lexical and grammatical nominalizations, let us examine the following data from the Austronesian language Malagasy, where, besides so-called focus morphology, which is a nominalizing morphology itself (see Section 3.2), there are several nominalizing prefixes that combine with different focus morphology. For example, in (2-23a) below, the nominalizer is a circumfix combining the f- prefix and the circumstantial focus suffix -ina (and its variants). The same circumfix is used in forming event grammatical nominalizations, as in (2-23b).

(2-23) Malagasy (Austronesian)\(^{14}\)
   a. Lexical nominalizations
      i. f-i-anar-ana
         NMLZR-AF.MM-advice-CF
         ‘school’
      ii. f-am-ono-na olona
         NMLZR-AF-kill-CF human
         ‘human killing, murder’
   b. Grammatical nominalizations
      i. mahafinaritra ny fandehehanana miaraka amin’ny
         mahafinaritra ny [f-an-dehadeha-nana miaraka amin-ny
         fun INDEF NMLZR-AF-walk(rdpl)-CF outside with-IND
         ankizy
         ankizy] child
         ‘Walking outside with children is fun.’
      ii. hitako ny fikapohana ilay alike
         hita-ko ny [f-i-kapoh-ana ilay alike]
         see.PF-1.PSG INDEF NMLZR-AF.MM-hit-CF DEF dog
         ‘The hitting of the dog was seen by me.’

\(^{14}\) The examples without mention of the sources are from the author’s own research.
As seen above, languages across the globe mark both lexical and grammatical nominalization similarly highlighting the underlying unity of the two types of nominalization. Many other languages, however, opt for marking lexical and grammatical nominalizations differently. This option is also motivated in that such a marking pattern reflects the differences in the types of nominalization, namely lexical nominalization vis-à-vis grammatical nominalization. We shall see in Section 8 how languages respond to the two opposing functional demands, one a desire to mark an underlying unity in the face of functional diversities, and the other a drive to mark functional differences of the forms at the expense of their underlying unity. Table 1 below summarizes the types of verbal-based nominalizations illustrated by some representative English forms, where the function-based classification cuts across the distinctions between lexical and grammatical nominalizations as well as formal morphosyntactic differences.

Table 1. Types of verbal-based nominalizations

<table>
<thead>
<tr>
<th>Lexical nominalizations</th>
<th>Event nominalizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(the) fight (lasted three minutes)</td>
<td>(the) employing (of Bill)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Argument nominalizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) cook, (a) judge, (an) alarm</td>
</tr>
<tr>
<td>employer</td>
</tr>
<tr>
<td>employee</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V-based nominalizations</th>
<th>Event nominalizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(that) John employs Bill</td>
<td>(John’s) employing Bill</td>
</tr>
<tr>
<td>to employ Bill</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grammatical nominalizations</th>
<th>Event nominalizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>what broke John</td>
<td></td>
</tr>
<tr>
<td>what John broke</td>
<td></td>
</tr>
<tr>
<td>(one) who employs Bill</td>
<td></td>
</tr>
<tr>
<td>(one) who(m) John employs</td>
<td></td>
</tr>
<tr>
<td>(one) which broke John</td>
<td></td>
</tr>
<tr>
<td>(one) which John broke</td>
<td></td>
</tr>
</tbody>
</table>
Our focus in this paper is on grammatical nominalizations, the treatment of which has far-reaching theoretical implications calling into question many descriptive practices and theoretical claims made in the field.

3. Verbal-based grammatical nominalization

Verbal-based grammatical nominalizations, like the ones given above, involve a verbal head possibly with nominal arguments and modifiers. We distinguish between event nominalizations and argument nominalizations. The former denote events and such abstract concepts as a state of affairs, a fact, or a proposition associated with the denoted events. They also denote event protagonists as well as resultant products, similar to resultative lexical nominalizations (e.g. *a building, a painting*). The argument nominalizations, on the other hand, denote in a clearer manner concrete things and thing-like entities, such as an agentive event protagonist, a patientive protagonist, a beneficiary, an instrument, or a location in close association with the concepts denoted by the verbal stems. That these two types of construction represent a unified phenomenon, namely grammatical nominalization, is indicated by a shared morphology in many languages such as the sharing of the particle *no* in the following Japanese examples, where the particle, glossed NPM (NP-use marker), marks a use of grammatical nominalizations as the head of a noun phrase—the *NP-use* of grammatical nominalizations (see Section 5).

(3-1) Event nominalization

Masako wa [otto ni sonna onna ga ita] no o sitta.

*Masako learned that (her) husband had such a woman.*

(3-2) Argument nominalization

i. [Asoko de Ø koi o tutte iru] no wa boku no otooto da.

*The one who is fishing a carp there is my younger brother.*

ii. [Asoko de otooto ga Ø tutte iru] no wa koi da.

*What my younger brother is fishing there is a carp.*

The entity-denoting function of grammatical nominalizations allow them to head an NP, the most telling syntactic property of nominals. In addition, they may function as a modifier in an NP, or they may play an adverbial function, where they denote such notions as simultaneously- or sequentially-occurring events, purpose,

15. These cover verb-based and adjective-based grammatical nominalizations in those languages in which adjectives also inflect as in Japanese.
reason, time, and location metonymically associated with events. These are all uses of grammatical nominalizations, not what grammatical nominalizations are per se, as we shall see below.

3.1 Event nominalizations

Event nominalizations have clause-like internal structures, often with a full array of NP arguments overtly expressed, as in (3-1) above. They have, however, external syntagmatic properties like nouns in that they head an NP, playing both syntactic and referential functions of arguments of a clause. Bear in mind that grammatical categories are determined on the basis of external properties, not by internal properties, meaning that even if a structure is clause-like internally, it does not follow that the structure in question is a clause (see Section 6). Event nominalizations denote the following kinds of concepts:

(3-3) a. Event/Activity

River Thompson Salish (Salish; Kroeber 1977)

\[cuk’w [e s-pekw’-e-s]_{NMLZ}\]
finish ART NMLZR-split-TRZ-3:TS
‘S/he finished splitting them.’

English

[John’s falling off of the bed]_{NMLZ} happened at 3:00 AM.
Cf. The event of [John’s falling off of the bed]_{NMLZ} happened at 3:00 AM.

b. Fact

Yaqui (Uto-Aztecan; Guerrero 2012)

\[Nim achaî [jaibu enchi siika-m]-ta te’a-k.\]
1SG.GEN father already 2SG.ACC go.SG.PFV-NMLZ-ACC find-PFV
‘My father discovered that you already left.’

Quechua (Cuzco dialect; Lefebvre & Muysken 1988)

\[Xwan papa-ta mikh-sqa-n]-ta yacha-ni.\]
Juan potato-ACC eat-NMLZ-3-ACC know-1SG
‘I know that Juan eats potatoes.’

c. Proposition

Northern Paiute (Western Numic; Uto-Aztecan; Thornes 2012)

\[ni [i = çadua-na] naka-supidakwatu.\]
1 2 = talk-NMLZ hear-understand
(lit.) ‘I understood your talking.’
‘I understood what you are saying.’

(Glosses modified and the literal translation supplied)

English

[That John is honest]_{NMLZ} is absurd.
Cf. The proposition [that John is honest]_{NMLZ} is absurd.
d. Event protagonist/participant
Navajo (Athabaskan; Hale & Platero 1973)
\[\text{Ashkii at’èéd yiyittså-(n)éé} \text{ yidloh.}\]
\[\text{boy} \text{ girl} \text{ saw-nmlzr laugh}\]
i. ‘The girl that the boy saw is laughing.’
ii. ‘The boy that saw the girl is laughing.’
Cf. \[\text{Ø ashkii yiyittså-(n)éé} \text{ at’èéd yidloh}\]
\[\text{boy} \text{ saw-nmlzr girl laugh}\]
‘The girl that saw the boy is laughing.’

e. Resultant product (“resultative nominalizations”)
Waiwai (Cariban; Derbyshire 1999: 57)
\[\text{a-mok-ɾɨ} \text{ 2-come-ac.nmlzr w-enta}\]
‘I heard you/your coming.’
Bolivian Quechua
\[\text{Maria laranjas-ta ch’irwa-sqa-n}-\text{ta uija-ni}\]
\[\text{oranges-acc squeeze-p.nmlz-3sg-acc drink-1sg}\]
(lit.) ‘I drink that Maria squeezing oranges.’ ‘I drank {juice resulting from} Maria’s squeezing oranges.’

f. Location
Gavião of Rondônia (Mondé; Tupi; Moore 2012)
\[\text{me-tá mât} \text{ ká téét mééy-ka} \text{ paágáá}\]
\[\text{2p-live nmlz.concrete in exact 2p-(aux.imperat.def)-go (3s)-open}\]
\[\text{kára-ále-á} \text{ yet-future-end}\]
‘Go open (it) where you live.’
Mosetén (Sakel 2004: 94, 95)
\[\text{chhiko’-ni-ti-dye’}\]
\[\text{liquid-put-VD-NMLZR}\]
‘place where one washes oneself’
cf. \[\text{saeks-e-dye’}\]
\[\text{eat-VI-NMLZR}\]
‘food’
\[\text{fer-dye’}\]
\[\text{strong-NMLZR}\]
‘strength’

g. Time
Mosetén (Sakel 2004: 95)
\[\text{iţts-i-dye-dyedye’-ra, añe-dye’}\]
\[\text{mature-VI-NMLZR-INC-FS-IR rain-NMLZR}\]
‘at the beginning of the time of the ripening, the rainy season (the time of rain)’
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

h. Reason
Musqueam (Suttles 2004: 105, 267)

\[ \text{š'íw} \; [k^{"w}\; [nə-šx^{"w}-ʔom\; 'í]] \]

important ART my-OBL-NMLZR-come

'I've come for an important reason.' (lit. 'The reason why I have come is important.')

Cf. \text{těčəθ} 'lie down' > šx^{"w}-těčəθ 'bed'

\text{yá'as} 'be working' > šx^{"w}-yá'as 'tools'

i. Manner
Měbengokre. (Je; Salanova 2011)

\[ [a-dju-jarẽnh] \]

\text{mex}

2-ANTIPASS-say.NMLZR good

'You spoke well.' (lit. 'Your saying was good.')

3.2 Argument nominalizations

When event nominalizations of the above type evoke event protagonists as denoted entities, potential ambiguity arises, as indicated by the Navajo example in (3-3d) above and the Slave example in (5-69b) below. Languages appear to cope with this problem in several ways.\(^1⁶\) The Siouan language Crow opts for marking the relevant argument by the indefinite specific determiner -\text{m} (Graczyk 2007: 262ff). In (3-4) below, the only possible reading is that it is “this one” who is concerned as having children, not “his sister-in-law”.

(3-4) \[ \text{hinne hawáta-}m\; úake \quad \text{dútt-ak\; áxpee-sh} \quad \text{kalakoon} \]

this one-DET his.sister-in-law takes-ss marry-DET then

\[ \text{dáak-uu-wish-dak} \quad \text{Graczyk 2007: 266} \]

child-PL-exist-COND

'if this one, who took his sister-in-law and married her, then has children.'

By far the more widespread method of unambiguously indicating the argument to be evoked is leaving the relevant argument position lexically unfilled, as in (3-2) above. The gap then indicates the grammatical role that the denotation of argument nominalizations stand for, such that a nominalization with a gap in subject position denotes an entity playing the subject role. These, paralleling the Japanese examples in (3-2), are illustrated by Korean subject argument nominalization (3-5a)

---

\(^{16}\) Word order within a nominalization structure seems to be relevant in Diegueño (Gorbet 1974).
They are grammatical counterparts of lexical argument nominalizations of the type, *employer* (agent nominalization) and *employee* (patient nominalization). The gap in these grammatical argument nominalizations functions as a variable that can refer to anyone matching the denotation, namely any individual fishing carp there for (3-5a) and anything that the speaker’s younger brother is fishing for (3-5b). The Korean pattern is paralleled by a large number of languages, in which the only clue for the type of argument nominalized is the position of a gap (or a missing argument) in the nominalization structure, as also seen in the Chinese examples below.

(3-5) Korean

a. [ceki-eyse Ø inge-lul nakk-ko iss-nun] key nay tongsaying-i-ya.
   there-at carp-acc fish-ger be-NMLZR NPM my y.brother-cop-ind
   ‘(The one) who is fishing carp there is my younger brother.’

b. [ceki-eyse nay tongsaying-i Ø nakk-ka-iss-nun] key inge-i-ya.
   there-at my y.brother-nom fish-ger-be-NMLZR NPM carp-cop-ind
   ‘What my younger brother is fishing there is carp.’

(3-6) Mandarin Chinese

a. [Ø zài nàr diào lìyú]=de shì Xiăo Wáng.
   PROG there fish carp=NMLZR COP Little Wang
   ‘(The one) who is fishing carp there is Little Wang.’

b. [Xiăo Wáng zài nàr diào Ø]=de shì lìyú.
   Little Wang PROG there fish carp=NMLZR COP carp
   ‘What Little Wang is fishing there is carp.’

Other languages depart from these patterns of argument nominalizations in two ways. One is the pattern found in a fair number of languages in the mainland Southeast Asia, Semitic languages, Iranian, and some others, where the argument nominalized is marked by a pronoun rather than by a gap. These languages actually combine the gap strategy and the pronoun strategy, where subject nominalization is marked by a gap, object nominalization by either a gap or a pronoun, and oblique nominalizations by a pronoun.18 Observe the following data from Thai,

17. Depending on the word order and the absence of case markers, argument nominalizations marked by a gap may still be ambiguous, as in Iskonawa discussed by Zariquiey (this volume). Zariquiey, however, points out an important difference between event nominalizations, which also evoke event protagonists, and argument nominalizations in that the former evoke only most prominent arguments (subject and object), while the latter may denote entities holding peripheral roles.

18. Kakataibo shows a pattern like this according to Roberto Zariquiey.
which may have a pronoun in subject position, and Modern Hebrew, where non-subject positions allow or require a pronoun.

(3-7) Thai (Yaowapat & Prasithrathsint 2009:7)

\[
\text{\textit{thể Maya khuan kin yaa [thii \ 0/man mòt?ayý?]}^19}
\]

2.sg not should eat medicine nmlzr \(0/3.sg\) expire

‘You should not take the medicine which expired.’

(3-8) Modern Hebrew (Semitic)

a. \(Zo\ \text{she} = [\ 0 \text{boxa}] \text{xavera sheli.}\)

\(\text{this.fem nmlzr cries is friend mine}\)

‘The (one) who is crying is my friend.’

b. \(Zo\ \text{she} = [\text{Yoav} \text{raa} \text{ota} \text{etmol}] \text{xavera sheli}\)

\(\text{this.fem nmlzr = Yoav saw her yesterday is friend my}\)

‘The (one) whom Yoav saw yesterday is a friend of mine.’

c. \(Ze [\text{sixakti ito} \text{etmol}] \text{haya shovav.}\)

\(\text{this.msc 1.played with him yesterday was naughty}\)

‘The (one) with whom I played yesterday was naughty.’

3.2.1 Role markers

While in Japanese, Thai, Hebrew, and many other languages, a gap or a pronoun is the only clue in ascertaining the nature of an argument evoked, others in addition have morphological markers that indicate the grammatical role of the evoked argument. The pattern that appears more widespread than others distinguishes subject (or agentive) and non-subject (or patientive/result) argument nominalization possibly with some additional distinctions, as in some dialects of Quechua (-q vs. -sqa), Turkish (-\(\text{-En}\) vs. -\(\text{-dIk/-cEk}\) + personal suffix), Tibeto-Burman Qiang(-\(\text{m}\) (and others) vs. -\(\text{O + gen}\)), and Yaqui, where, as seen below, the argument positions nominalized are represented by a gap for central grammatical relations and with person-marked relational particles, e.g. a-\(\text{mak}\) in (3-9d), for peripheral relations with additional morphology (-\(\text{me}\) for subject, -\(\text{\prime u}\) for non-subject, and -\(\text{\prime Vpo}\) for locative) indicating the grammatical roles that are nominalized.

\[\text{19. My Thai consultant finds the form with the pronoun in subject position less felicitous than the one with a gap. Riddle (1993: 60) gives White Hmong as another language that permits a “resumptive pronoun” in subject position of a relative clause. A Kwa language in Ghana, is a rare type requiring a pronoun in the subject position of argument nominalizations as well (Campbell 2017).}\]
(3-9) Yaqui (Southern Uto-Aztecan; Albert Alvarez 2012)

Subject nominalizer: -me.

a. Jabesa [Ø wa-me yabe-m tea-ka-me]
   who  DEM-PL key-PL find-PERF-S.NMLZR
   ‘Who is the one that found the keys?’

Non-subject nominalizer: -'u.

b. Jitasa [Joan-ta Ø tea-ka-'u]
   what  John-GEN find-PERF-NS.NMLZR
   ‘What is the one that John found?’

c. [em rebo'osam ameu jinu-ka-'u]-m
   2SG.GEN mantilla 3PL.DAT buy-PERF-NS.NMLZR-PL
   ‘ones from whom you bought mantilla’

d. [nim a-mak yepsa-ka-'u]
   1SG.GEN 3SG.ACC-COM come-PERF-NS.NLZR
   ‘one with whom I came’.

Locative nominalizer: -'Vpo

e. [jamuch-im Ø to'e'-epo]
   woman-PL sleep-L.NMLZR
   ‘(place) where women are sleeping’.

Far more complex patterns are seen in South America. The Carib language Hixkaryana has markers of event nominalizations distinct from those marking the role types of argument nominalizations, both of which have variant forms marking tense as well.

(3-10) Hixkaryana (Carib; Derbyshire 1999: 48–49)

a. Event/Action nominalizer: -(ni)ri
   a-wanota-niri
   2-sing-AC.NMLZR
   ‘your singing’

b. Event/Action nominalizer-Past tense: -thiri
   i-wanota-thiri komo
   3-sing-AC.NMLZR COLL
   ‘their singing (in the past)’

c. Nominalizer of A (Agentive protagonist of transitive event): -ne
   co-hananih-ne
   1-instruct-A.NMLZR
   ‘one who instructs me’

d. Nominalizer of the S (Protagonist of intransitive event)/O (Patientive protagonist of transitive event)-Past tense: -saho
   S: i-manho-saho urotomoro/moki
   IMPERS-dance-S.NMLZR 1/2/3PRO
   ‘I (am)/you (are)/he (is) the one who danced.’
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

O: \( t\text{-}ono\text{-}saho \) koso
   IMPERS\text{-}eat\text{-}O.NMLZR deer
   ‘The deer (was) what was eaten.’

e. Nominalizer of thing/time/place associated with State/Action: \( t(o)ho \)
   r\text{-}omoh\text{-}toho
   I\text{-}come\text{-}ASSOC.NMLZR
   ‘the time of my coming’/‘place to which I’ll come’/‘circumstance of my coming’
   \( i\text{-}hkoto\text{-}tho \)
   3\text{-}cut\text{-}ASSOC.NMLZR
   ‘saw, thing for cutting it’

f. Nominalizer of Negation: \( -hini/-hni \)
   i\text{-}to\text{-}hni \( u\text{co} \)
   IMPERS\text{-}go\text{-}NEG.NMLZR 1PRO
   ‘I (am) one who does not/will not go.’

While in Hixkaryana O and S are treated alike showing an ergative pattern in
nominalization (3-10d), the Tupian language Kamaiurá distinguishes not only be-
tween O and S, but also between O (patient) and P (theme).

(3-11) Kamaiurá (Tupí-Guarani; Seki 2000 and p.c.)\(^{20}\)

A: \( -tat/-tar \)
a. \([O\ ka’i\text{-}a\ juka\text{-}tar\text{-}er\text{-}a]\)
   monkey\text{-}NU kill\text{-}A.NMLZR\text{-}PST\text{-}NU
   ‘(one) who killed a monkey’

S: \( -ma’e \)
b. \( o’\text{-}ata\text{-}ma’e \) \( \text{(122)} \)
   3\text{-}walk\text{-}S.NMLZR
   ‘that which/who walk’

c. \( i\text{-}pitsun\text{-}ama’e \) \( \text{(179)} \)
   3\text{-}black\text{-}S.NMLZR
   ‘that which is black’

O (patient): \( -emi \)
d. \( j e=r=emi\text{-}juka \)
   LSG\text{-}RELATIONAL\text{-}O.NMZ\text{-}kill
   ‘what I killed’

---

\(^{20}\) Apparently the functions of the cognates of these markers are somewhat different in other
Tupí-Guaraní languages (see Jensen 1999: 160).
P (theme): -ipy/-pyr
e.  i-mono-pyr-er-a  Sau Paulo katy
   3-send-P.NMZR-PST-NU direction
   ‘one sent to Saõ Paulo’
Location/Instrument: -tap/-ap/-taw/-aw
f.  [i-jo-taw-er-a]
   3-go-L.NMLZR-PST-N
   ‘place where he went out from’
g.  [moī-a  juka –taw-er-a]
   cobra-N kill-L.NMLZR-PST-N
   ‘thing with which I killed the snake’

The original function of the so-called Austronesian focus-system appears similar
to these role markers of argument nominalizations (Starosta, Pawley, and Reid
1982). The AF (actor-focus) form marks subject/agent argument nominalizations,
Pf (patient-focus) form object/patient argument nominalizations, LF (locative-
focus) form locative argument nominalizations, and CF (circumstantial-focus)
form beneficiary and instrumental argument nominalizations. Many Austronesian
languages in Taiwan, Malaysia, and Indonesia have reduced the proto-Austrone-
sian four-way focus pattern to a three- or a two-way contrast, but many others in
Taiwan, the Philippines, as well as Malagasy largely have reflexes of the original
four-way contrast, as seen in the following forms from Mayrinax Atayal in Taiwan.

(3-12)  Mayrinax Atayal (Austronesian; based on Huang 1995)
a.  Actor focus  (Actor nominalization)
   ūaq-un=mu  ku?  [m-aqwas Ø]  kaʔ hacaʔ
   know-PF=1SG.GEN NOM.REF  AF-sing  LIN that
   ‘I know that singer/one who is singing.’
b.  Patient focus  (Patient nominalization)
   ma-hnuq ku?  [β-in-ainay  Ø nukʔ  naʔakash]
   AF-cheap NOM.REF  buy<PF.REALIS>buy  GEN.REF old.man
   ‘What the old man bought was cheap.’
c.  Locative focus  (Locative nominalization)
   yahapuyan ku?  [naniq-an Ø cuʔ  βuŋaʔ  nkuʔ  rulaqiʔ]
   kitchen NOM.REF  eat-LF  ACC.NONREF yam  GEN.REF child
   ‘The kitchen is (the place) where the child eats yam.’
d.  Circumstantial focus  (Benefactive/instrumental nominalization)
   ini=mu  swaʔ=I kuʔ  [si=ghahapuy Ø nkuʔ  kanairil]
   NEG=1SG like=LF NOM.REF  CF=COOK  GEN.REF woman
   ‘I don’t like the one for whom the woman cooks.’
Turning to European languages, German shows a very systematic pattern similar to the Austronesian focus system, but in terms of demonstrative-based nominalization markers. Observe:

(3-13) German
a. *Ich empfange den, [der] [Ø morgen kommt].*
   'I receive the one who comes tomorrow.'

b. *Ich empfange den, [den] [du mir Ø vorgestellt hast].*
   'I receive the one whom you introduced to me.'

c. *Ich empfange den, [dem] [du Ø den Brief gegeben hast].*
   'I receive the one whom you gave the letter.'

d. *Ich empfange den, [dessen] [Ø Buch ich gelesen habe].*
   'I receive the one whose book I have read.'

Modern English has partially lost the marking distinction between subject and object nominalization, but the marker *whom* still uniquely marks an object nominalization. It also has distinct markers for adverbial nominalizations, denoting a place, a time, etc., as below, where we are reanalyzing so-called relative pronouns as nominalizers.

(3-14) English
a. *You should marry [who [Ø loves you]].* (SU nominalization)

b. *You should marry [who/whom [you love Ø]].* (OBJ nominalization)

c. *The man [whose [Ø book has just been published]] is in town.* (GEN nominalization)

d. *The book hits [where [it hurts Ø most]].* (Place nominalization)

e. *That was [when [I ruled the world Ø]].* (Time nominalization)

f. *[Why [he didn’t come to the party Ø]] remains a mystery.* (Reason nominalization)

g. *You might not like [how [he eats his food Ø]].* (Manner nominalization)

English also distinguishes forms for animate and inanimate denotations in terms of *who(m)* and *which/what.*

21. It is not clear if we have a gap in this and the structures in (f) and (g).
(3-15) English
a. You may choose [who(m) [you like Ø]]. (Animate OBJ nominalization)
b. You may choose [which/what [you like Ø]]. (Inanimate OBJ nominalization)

The classificatory function of nominalizers is observed in Newar and many other languages that have nominalizing classifiers (see Section 8).

(3-16) Newar (Tibeto-Burman; courtesy of Kazuyuki Kiryu)
a. [ana Ø dan-ā cwā:=mha] rām=yā macā kha:. there stand-cm exist.ND=NMLZR Ram=GEN child COP
(Animate SUB nominalization)
‘The one standing there is Ram’s child.’
b. [ana Ø du=gu] rām=yā gāri kha:.
there exist.ND=NMLZR Ram=GEN car COP
(Inanimate SUB nominalization)
‘The one that is there is Ram’s car.’

These classifying nominalizations play a vital role in our understanding of nominalizations as nominal structures that denote substantive entities classifiable according to features such as gender, animacy, physical shape, and function. The overall importance of the role markers discussed above in the formation of relative clause constructions is discussed in Section 9.2.

4. Formal representations of verbal-based nominalizations

We assume that nouns have a denotation index that connects a noun with a set of substantive concepts it denotes, as in the following manner, where the arbitrary numbers are indices that connect nouns (forms) and the entity concepts that they denote (meanings).

(4-1) Noun CONCEPT

a. [dog]N

b. [cat]N

c. [flower]N

Nominalization is a process that creates linguistic forms (words and complex structures) that have form-concept pairings similar to nouns, as below.
Lexical nominalizations (4-2a) and (b) are like ordinary nouns, except the former is derived via zero morphology, and are straightforward. While nouns have permanent form-concept relationships registered in the speaker’s mind, grammatical nominalizations establish form-concept relationships for the nonce, whereby nominalization structures as a whole bear denotation indices. In event nominalization (4-2c), the entire structure may denote a kind of fact, namely the one associated with the event of John’s coming yesterday. In the case of argument nominalizations, nominalization structures bind an argument position that is empty, as in the English examples or that may be occupied by a pronoun, as in the Hebrew example (4-2f). This binding means that what the nominalization structure denotes plays the grammatical role of co-indexed NP position. Subject nominalization (4-2d), for example, denotes a substantive entity that plays a subject role; e.g. one who loves the addressee. Object nominalizations (4-2e), (f) denote substantive entities that play an object role; e.g. one the addressee loves or one Yoav saw yesterday.

Unlike the usual analysis of relative clauses that assumes so-called relative pronouns to bind an empty argument position, our analysis makes the entire nominalization structures bear a denotation index, just like ordinary nouns, which directly binds an empty argument/adjunct position, or one filled by a pronoun, both of which function as a variable. In this analysis so-called relative pronouns play no pronominal role; they are treated as nominalizers whose functions are similar to nominalizing morphology or nominalizing particles in other languages. This is a desired result since many languages of the world, as we shall see presently, do...
not involve any pronouns, like the so-called relative pronouns in English, in their formation of relative clause constructions.

As in the other cases of metonymy, argument nominalizations evoke a variety of denotations and context determines the appropriate denotations per the Gricean Cooperative Principle. For example, Spanish subject nominalization (4-3) below denotes a variety of masculine entities that are white, such as a white stallion, a white male dog, a white hat, and a white car. In the context of (4-4), the denotation of \{WHITE HAT\} will be chosen, while in the context of (4-5), the denotation of \{WHITE CAR\} will be selected.

\[(4-3) \quad [\text{que} \quad [\emptyset \text{ es blanco}] ]
\]
\[\text{nmlzr} \quad \text{is white}
\]

\[\text{`that which is white'}
\]

(4-4) a. ¿Qué tipo de sombrero te gusta?
What kind of hat do you like.3sg

‘What kind of hat do you like?’

b. Me gusta uno [que [\emptyset \text{ es blanco}]]
me like.3sg one \text{nmlzr} \quad \text{is white.}

‘I like one which is white.’

(4-5) a. ¿Qué tipo de coche te gusta?
what kind of car do you like.3sg

‘What kind of car do you like?’

b. Prefiero uno [que [\emptyset \text{ es blanco}]].
prefer.1sg one \text{nmlzr} \quad \text{is white}

‘I prefer one which is white.’

Having elaborated on the nature of grammatical nominalizations, we shall now turn to their use and function.

5. Structures and their use

One of the major problems we find in some current definitions of nominalization in dictionaries and encyclopedias, and even in professional writings is the confusion over a form and its use. For example, Crystal’s definition discussed in the introduction says that nominalization may involve “the derivation of a noun phrase from an underlying clause” (Crystal 1980: 328). In a similar vein, Givón (2009: 66) defines nominalization as “the process via which a finite verbal clause – either in its entirety or only the subject-less verb phrase – is converted to a noun phrase.”
Just as ordinary nouns are different from noun phrases, nominalizations are different from noun phrases, though both may be used as the head of a noun phrase. Nouns and nominalizations may also be used as a modifier that may restrict or identify the denotation of a head noun. These two major uses of nominals are illustrated below, first using an ordinary noun.

(5-1) NP-use/Referential function

\[ [\text{Cotton}]_N \text{NP} \text{ is a useful commodity.} \]

(5-2) Modification-use/Restrictive function

a. \[ [[\text{cotton}]_N \{\text{mill}\}_N]_N \] (noun compound)
   cf. lumber mill
b. \[ [[\text{cotton}]_N, \{\text{shirt}\}_N]_\text{NP} \] (noun phrase)
   cf. silk shirt

In (5-1) the noun \([\text{cotton}]_N\) heads an NP and has a referential function at the NP level, referring to a type of commodity in the real world. We distinguish between “denotation” and “reference”. The former is the relation between words/structures and their associated mental concepts, as described in the preceding section, while the latter is the relationship between noun phrases headed by words/structures and entities in the real (or imaginary) world. Thus nouns and nominalizations themselves have denotations but they do not directly refer. When they head an NP, as in (5-1), they play a referential function as the main constituent of an NP. Reference is a type of speech act of pointing out a real world entity by the use of a nominal form as the head of nominal phrase of N or NP upon recognition of a real world entity in question as an instance of the denotation of the nominal form being used.\(^{22}\) In the modification-use, as in (5-2), nouns do not refer. Instead, their denotations function to restrict the denotation of a head nominal. Notice that these modifying nouns remain nouns; in particular, they do not become adjective even when they play a modification function. This is shown by the fact that a modifying noun takes an adjectival modifier, not an adverbial modifier that adjectives would take. Observe:

(5-3) a. \[ [[\text{Egyptian cotton}]_N \{\text{shirt}\}_N]_\text{NP} \]
b. \[ [[\text{new/newly car}]_N \{\text{smell}\}_N]_\text{NP} \]

Figure 1 shows the relationship between a structure (noun in our case) and its use and the functions associated with the use.

---

\(^{22}\) Cf. Searle (1969: 85): “[reference is] an act of identifying some entity that the speaker intends to talk about.”
Masayoshi Shibatani

Structure Use/Function

NP-use/Referring function

[[Cotton]N]NP is a useful commodity.

[cotton]N

Modification-use/Restrictive function


cf. [[silk]N [shirt]N]NP

Figure 1. Two uses of noun

Grammatical nominalizations, qua quasi-nominals, behave like ordinary nouns in allowing both NP- and modification-use, as below:

(5-4) NP-use of event nominalization
a. \([[That John is honest]_{nmlz}]_{NP} \) is absurd.
b. \(I \ know \ [[that \ John \ is \ honest]_{nmlz}]_{NP}\)

(5-5) Modification-use of event nominalization
a. \([[The \ proposition \ [that \ John \ is \ honest]_{nmlz}]_{NP} \) is absurd.
b. \(I \ know \ [the \ fact \ [that \ John \ is \ honest]_{nmlz}]_{NP}\)

(5-6) NP-use of argument nominalization
a. \([[Who \ [Ø gets \ there \ first]]_{nmlz}]_{NP} \) defines the truth. (David Baldacci)
b. \(You \ should \ marry \ [[who \ [you \ love \ Ø]]_{nmlz}]_{NP}\)

(5-7) Modification-use of argument nominalization
a. \([[The \ man \ [who \ [Ø gets \ there \ first]]_{nmlz}]_{NP} \) defines the truth.
b. \(You \ should \ marry \ [a \ man \ [who \ [you \ love \ Ø]]_{nmlz}]_{NP}\)

The two uses of an argument nominalization structure are diagrammatically shown in Figure 2 (next page).

5.1 NP-use and modification-use of grammatical nominalizations

As demonstrated by many of the contributions to this volume, these patterns of use of grammatical nominalizations are not unique to English or European languages.
Just to drive this point home, we shall first provide random samples from genetically diverse languages of the Americas and elsewhere illustrating the NP- and modification-use of argument nominalizations paralleling the patterns of (5-6) and (5-7) above, whose far-reaching theoretical significance will be discussed in the subsequent sections. The examples labeled (a) represent the NP-use and those labeled (b) the modification-use of argument nominalizations.

(5-8) Toba (Guaicuruan; Messineo & Porta 2009 and Cristina Messineo p.c.)
   a. \[\text{na} [\text{Ø-chigochigina yi Espinillo}_{\text{NMLZ}}]_{\text{NP}} \text{Ø-tayge da Salta ko’ollaGa} \]
      \[\text{DD 3-come from DD Espinillo 3-go DD Salta PST} \]
      ‘Those who came from the Espinillo went to Salta.’
   b. \[\text{ajem si-kjen} [\text{so fijaGawa} [\text{Ø-neta-ge da cako}_{\text{NMLZ}}]_{\text{N}}]_{\text{NP}} \]
      I 1A-greet DET man 3-be-DIR DET Chaco
      ‘I greeted the man who lives in Chaco.’

(5-9) Tapiete (Tupí-Guaraní; Ciccone 2008)
   a. \[\text{á-ha-po a-hapi} [(\text{kwé(we) a-yasiya-wa})_{\text{NMLZ}}]_{\text{NP}} \]
      1SG.AC-go-FUT 1SG.AC-light before 1SG.AC-cut-NMLZR
      ‘I am going to light what I cut last time.’
   b. \[\text{hau ye waka ro’o a-}m bầu-\text{wu} [\text{NMLZ}]_{\text{NP}} \]
      1:eat already cow meat 1SG.AC-heat-NMLZR
      ‘I already ate the meat that I heated.’

(5-10) Kipeá (Macro-Jê; Rodrigues 1999: 104, 195)
   a. \[[\text{di-te-ri}}_{\text{NMLZ}}]_{\text{NP}} \]
      ABS-come-NMLZR
      ‘the one who comes’
   b. \[[\text{w}a\text{re du-di-ri udza}}_{\text{NMLZ}}\]_{\text{NP}}
      priest ERG-give-NMLR knife
      ‘the priest who gave a knife’
(5-11) Kakataibo (Panoan; Zariquiey 2011: 634, 642)
   a. *Marianën ʼakukë a kana pian*
      [Maria-nën ʼaku-kë] NMLZ a kana pi-a-n
      Maria-ERG cook-NMLZR 3SG.O NAR.SG eat-PERF1/2P
      ‘I ate what Maria cooked.’
   b. *ain bënên ʼakë buë*
      [ain bënë=n ʼa-kë] NMLZ buë
      3.GEN husband=ERG do-NMLZR fish.esp
      ‘the fish that her husband fished’

(5-12) Bolivian Quechua (Quechuan)
   a. *[[Maria-q waykʼu-sqa-n] -ta mikʼu-sayku]*
      Maria-GEN cook-O.NMLZR-3SG-ACC eat-PROG.1PL.EXCL
      ‘We are eating what Maria cooked.’
   b. *[[Maria-q waykʼu-sqa-n] wallpa]-ta mikʼu-sayku*
      Maria-GEN cook-O.NMLZR-3SG chicken-ACC eat-PROG.1PL.EXCL
      ‘We are eating the chicken that Maria cooked.’

(5-13) Hixkaryana (Carib; Derbyshire 1999: 48, 57)
   a. NP-use of O-nominalization (-saho)
      *[t-ono-saho] NMLZ koso*
      IMPERS-eat-O.NMLZR deer
      ‘The deer (was) what was eaten.’
   b. Modification-use of O-nominalization (-ni)
      *miriri erahma-phi-i-ya ti-rui ni-kupi-hipi*
      that see-PST-3-ERG 3REFL-brother O.NMLZR-do-PST
      ‘He saw that, what his brother had done.’

(5-14) Tuyuca (Tucano; courtesy of Janet Barnes)
   a. *[[baco-á-ri-gi]]*
      to.have.been.bitten-recent-SG.NMLZR-CLF
      (cylindrical.shape,long.and.solid)
      ‘that which (=a cylindrically-shaped, long and solid thing) was bitten’
   b. *[[niká [baco-á-ri-gi]]]*
      leg to.have.been.bitten-recent-SG.NMLZR-CLF
      ‘the leg that was bitten’

(5-15) Piapoco (Arawak; Klumpp & Burquest 1983: 390, 395)
   a. *yà-a-wa [i-té-eyéi-ca yà-ana]*
      3 m-go-aspect 3 m-carry-[+pl]-aspect 3 m-limb
      ‘(the ones) who carry the animal’s leg go’
   b. *niái inanaíà [i-yamé-eyéi-cawa capii iricu]*
      those women 3-stay-[+pl]-aspect house in
      ‘those women who stayed in the house’
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(5-16)  K’ichee’ (Mayan (Quichean); Larsen & Norman 1979: 357 and courtesy of Nora England and Telma Can Pixabaj)

a. $x$-$\emptyset$-\text{inw} \quad \text{\{lee \ [x-$\emptyset$-\text{u-}\text{ch’ay \ lee}
\text{ASP-3SG.ABS-1SG.ERG-see} \ \text{NMLZR} \ \text{ASP-3SG.ABS-3SG.ERG-hit the}
\text{achih}\}_{\text{NMLZ}}\}}_{\text{NP}}$\text{ man}

'I saw the one whom the man hit.' (Or 'I saw the one who hit the man.')</n
b. \text{\{leeixoq \ [lee \ [x-$\emptyset$-\text{u-}\text{ch’ay \ lee \ achih]}\}_{\text{NMLZ}}\}}_{\text{NP}}$

'the woman NMLZR ASP-3SG.ABS-3SG.ERG-hit the man

'the woman whom the man hit' (Or 'the woman who hit the man')

(5-17) Yucatec (Mayan (Yucatecan); Gutiérrez-Bravo 2012: 262, 264)

a. $\text{Yaan-}$-$\emptyset$ \quad \text{\{k-u \ wéej taal bejla\}}_{\text{NMLZ-e'}}$

ex-abs.3SG HAB-ERG.3SG still come today-CL

'There are those that still come today.'

b. $\text{Tuláakal le gente \ [k-u \ taal]}_{\text{NMLZ-o'}}$, $k$-$u$

all DM people HAB-ERG.3SG come-CL HAB-3.ERG.3SG
give-ind-ABS.3SG ERG food

'All the people that came, he gave them their food.'

(5-18) Nieves Mixtec (Otomamguean; Caponigro, Torrence & Cisneros 2013: 70, 76)

a. $\text{[yō$^{23}$ ni-kānī jēráldó] ni-kānī jwán}$

NMLZR CMP-hit Geraldo CMP-hit Juan

'The one(s) who hit Geralido hit Juan too.'

b. $\text{[jwán kūt óó = ra } \text{ñáa} [\text{yō kūt óó jēráldó}]$

Juan like.con = 3SG.M woman NMLZR like.con Geraldo

'Juan likes the woman who Geralido likes.'

(5-19) Yaqui (Southern Uto-Aztecan; Alvarez 2011)

a. $\text{[Junu’u weyeka-me]}$ \text{nakapit}$

dem be.standing-sub.NMLZR deaf

'This one who is standing is deaf.'

b. $\text{[Joan uka chu’u-ta [Maria-ta ke’e-ka-m]-ta me’a-k}$

John det.acc dog-acc Mary-acc bite-perf-sub.NMLZR-acc kill-perf

'John killed the dog that bit Mary.'

\footnote{23. The original gloss was “who”, which, like English wh-nominalizations/relatives, is treated as a nominalizer in this paper. See Mithun (2012) for more examples illustrating the use of interrogative pronouns as nominalizers.}
(5-20) Tümpisa (Panamint) Shoshone (Northern Uto-Aztecan; Dayley 1989: 476)
   a. [ke *tamangkan-tü*]$_{nmrz}$ *naamaa setü*. [Nümni *appü*] *utü*.
      not tooth.having-NMLZR was this our(excl) father that
      ‘He is one who is missing a tooth. That is our father.’
   b. *Tangumü* [niu *pusikwa-tü*]$_{nmrz}$ *tühüyanna kuttihantü*.
      man me know-NMLZR deer.o shot-stv
      ‘The man who knows me is shooting the deer.’

(5-21) Navajo (Southern Athabaskan; Willie 1989: 415, 435)
   a. [*’at’eéd yizts’q-s-yéé*]$_{yaltí’}$
      girl 3sO: 3sS:kiss-NMLZR speaking
      ‘The one who kissed the girl is speaking.’
   b. [*’ashikii [*’at’eéd yizts’q-s-yéé*]]
      boy girl 3sO: 3sS:kiss-NMLZR
      ‘the boy who kissed the girl’

(5-22) Lakhota (Siouan; Van Valin 1977: 81)
   a. [[*su’kawakhâ wa ima’kicu ki* he]$_{nmrz}$]$_{NP}$ *wâyá’ke*
      horse a he.takes.it.from.me DET that he.sees.him
      ‘He saw the one who took a horse from me.’
   b. [[*wicha’s*]$_{ânhnika}$ [*su’kawakhâ wa ima’kicu ki* he]$_{nmrz}$]$_{NP}$
      man horse a he.takes.it.from.me DET that
      he.sees.him Q
      ‘Did he see the man who took a horse from me?’

(5-23) Cherokee (Iroquoian; Montgomery-Anderson 2008: 523, 560)
   a. *júúskwakahli jituútőoʔa*
      [ti-uu-x skwakahli ji-tee-uu-at- ʔoʔa]
      dst2-3b-striped NMLZR-dst-3b-be.called:prc
      ‘what is called “Striped”’
   b. *askaya jijijiyalii nohehtiiskv*$_{aânhnika}$
      a-skaya [ji-jii-ali-hnohehtski-vrir]$_{a-ahnika}$
      3A-man NMLZR-1A-AN-MDL-talk.with:INC-EXP\SUB 3A-leave:IMM
      ‘The man that I was talking to left.’

(5-24) Creek (Muskogee) (Muskogean; Martin 2011: 391, 394)
   a. [[ilki-acól-i lêyk-a:t]$_{l-ólo:y-atís}$]
      3pat.father-old-1 sîl.sg.fgr-ref dir-reach.du.lgr-pst5-ind
      ‘They got to [where their elderly father lived].’
   b. asêy ifâ [a:hôy-t-a:t]$_{pô-pêy-i-t ô-s}$
      that dog dir.stand.fgr-ref nice-dur-t be.fgr-ind
      ‘That dog standing over there is friendly.’
(5-25) Rainy River Ojibwa (Algonquian; Johns 1982: 161, 162)

a. \[ka^{24}:-nagama-t \]kino:zi
   \[nmlzr\text{-sing}-3c \text{ tall.3}\]
   ‘The one who is singing is tall.’

b. \[iniini \{ka-\text{-nagama-t}\} \] kino:zi
   man \[nmlzr\text{-sing}-3c \text{ tall.3}\]
   ‘The man who is singing is tall.’

(5-26) St’át’imcets (Lillooet) (Salish; Davis 2010: 11)

a. \[ats’x-en=lhkácw=ha \] i=[\(t’iq\)=a]
   see-\text{TR}=2sg.SU=YNQ PL.DET=\{[\text{arrive}]=\text{EXIS}\}
   ‘Did you see [the ones who arrived]?’

b. \[ats’x-en=lhkácw=ha \] i= \{\(t’iq\)=a\} \[sqáyqeycw\]
   see-\text{TR}=2sg.SU=YNQ PL.DET=\{[\text{arrive}]=\text{EXIS men}\}
   ‘Did you see [the men who arrived]?’

(5-27) Central Alaskan Yupik (Eskimo-Aleut; Miyaoka 2012: 533, 543)

a. \[\text{[that-ee.abs.sg fish-rel.sg eye-rel.3sg.sg eat-a.nmlzr-abs.3sg.sg kass’a-u-lími-uq]}
   \text{white.man-be-evd-ind.3sg}\]
   ‘(I see now) that the one who is eating the fish eye is a white man.’

b. \[\text{[naq-e-m nere-sti-i]} \] qimugta
   fish-rel.sg eat-a.nmlzr-abs.3sg.sg dog.abs.sg
   ‘the dog that eats the fish’

Just to round things out, let us observe similar uses of argument nominalizations elsewhere around the globe, starting in the East Asia.

(5-28) Korean

a. \[\text{[Yenghi-ka ilk-un]} \] kes-un acwu elype-ta.
   Yonghee-NOM read-NMLZR NPM-TOP very difficult-IND
   ‘What Yonghee read is very difficult.’

b. \[\text{[Yenghi-ka ilk-un]} \] chayk-un acwu elype-ta.
   Yonghee-NOM read-NMLZR book-TOP very difficult-IND
   ‘The book that Yonghee read is very difficult.’

(5-29) Mongolian (Chakhar, Kesigten subdialect); courtesy of Bayaerduleng and Benjamin Brosig)

a. \[\text{[t’snt tfjg53-tf pae-gai]} \] nmlz=\text{n}\]
   man-\ae: xu:xat
   there stand-CVB be-PRS.NMLZR=3POSS we-gen child
   ‘The one standing there is our child.’

24. The original gloss for this was “WH-.” See footnote 19.
b. \[[t^{\text{ont} \text{fgs}^{\text{t}}} \text{pae}^{\text{-} \text{cat}}]_{\text{nmlz}} \text{xuxax}]_{\text{NP}} \text{p} \text{ol man-æ: there stand-cvb be-prs.nmlzr child} \text{TOP we-gen}

'The child who is standing there is ours.'

(5-30) Mandarin Chinese

a. \[[Wô zuótiān mài-de]_{\text{nmlz}}]_{\text{NP}} hên guèi.

I yesterday buy-nmlzr very expensive

'What I bought yesterday was very expensive.'

b. \[[Wô zuótiān mài-de]_{\text{nmlz}} shū]_{\text{NP}} hên guèi.

I yesterday buy-nmlzr book very expensive

'The book that I bought yesterday was very expensive.'

(5-31) Thai (courtesy of Kingkarn Thepkanjana)

a. chán chɔ̂ɔ p \[[thîi khwêen nay tûu]_{\text{nmlz}}\]_{NP}

I like nmlzr hang in closet

'I like the one that is hanging in the closet.'

b. chán chɔ̂ɔ p \[kràprooŋ thîi khwêen nay tûu]_{\text{nmlz}}\]_{NP}

I like skirt nmlzr hang in closet

'I like the skirt that is hanging in the closet.'

(5-32) Mayrinax Atayal (Austronesian; based on Huang 1995)

a. ßaq-un=mu ku? \[[m-aquwas]_{\text{nmlz}} ka? haca?]\]_{NP}

know-PF=1sg.gen nom.ref AF-sing lin that

'I know that one who is singing there.'/ 'I know that singer there.'

b. ßaq-un=mu ku? \[kanairil ka? [m-aquwas]_{\text{nmlz}} ka? haca?]\]_{NP}

know-PF=1sg.gen nom.ref woman lin AF-sing lin that

'I know that woman who is singing there.'

(5-33) Standard Indonesian (Austronesian)

a. Aku makan \[[yang diberikan ayah kepada-ku]_{\text{nmlzr}}\]_{NP}

I eat nmlzr give father to-me

'I ate what father gave to me.'

b. Aku makan \[ikan [yang diberikan ayah kepada-ku]_{\text{nmlzr}}\]_{NP}

I eat fish nmlzr give father to-me

'I ate the fish that father gave to me.'

(5-34) Sasak (Pancor ngeno=ngené dialect; Lombok Island, Indonesia; Austronesian)

a. Beng oku \[[si léq méje]_{\text{nmlz}} ino]\]_{NP}

give I nmlzr on table the

'Give me the one that is on the table.'

b. Beng oku \[buku [si léq méje]_{\text{nmlz}} ino]\]_{NP}

give I book nmlzr on table the

'Give me the book that is on the table.'
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(5-35) Kalkatungu25 (Pama-Nyungan; Australia; Blake 1979: 101)

a. kaanta-nə pakaik-ka kalpuru-tiŋu [(niŋ-ti ŋu-ŋa ŋapa)nmlz]NP
   leave-pst that-Ø Boulia-ABL you-ERG NMLZR-ACC saw
   ‘The one whom you saw left Boulia.’ (Ø = a morpheme without a
   referential content)

b. ŋai utantiji-nə [pa-u ɣaur-ku [nin-ti ŋu-ŋa laji]nmlz]NP
   I look after-pst that-DAT child-DAT you-ERG NMLZR-ACC hit
   ‘I’ve been looking after that kid you belted.’

(5-36) Telugu (Dravidian, southern India; courtesy of K. V. Subbarao)

a. neenu [[John icc-in-a]nmlz]NP di26 cadiveenu
   I give-pst-NMLZR NPM read
   ‘I am reading what John gave me.’

b. neenu [[John icc-in-a]nmlz pustakam]NP cadiveenu
   I give-pst-NMLZR book read
   ‘I am reading the book that John gave me.’

(5-37) Hindi (Indo-Aryan; courtesy of Miki Nishioka)

a. main [us kii xariidii huii] paRh raha
   I s/he.OBL GEN.F.SG buy.PFV.F SG/PL be.PFV.F SG/PL read PROG.M.SG
   hU
   be.PRS.1SG
   ‘I am reading what s/he bought.’

b. main [us kii xariidii huii] kitaab] paRh raha
   I s/he.OBL GEN.F.SG buy.PFV.F SG/PL be.PFV.F SG/PL book.F.SG read
   hU
   PROG.M.SG be.PRS.1SG
   ‘I am reading the book that s/he bought.’

(5-38) Abkhaz (North West Caucasian; courtesy of George Hewitt)

a. [[‘jy.b.taxy.w]nmlz]NP d ga
   whom.you(FEM).want.Non-Finite/STAT/PRS 3SG.take(IMP)
   ‘Take whom you (Female) want!’

b. [[‘jy.b.taxy.w]nmlz a-xa’ə]NP
   whom.you(FEM).want.Non-Finite/STAT/PRS article-man
   d-aa-wëit’
   he-comes-PRS/Finite/Non-STAT
   ‘Here comes the man whom you want.’

25. Kalkatungu being an extremely “flat” language, the constituency of relevant phrases below is
   not entirely certain. It is clear, though, that there is a nominalizer and that a nominalization can
   function as a subject and can modify a noun.

26. See the discussion on this particle labeled npm (NP-use marker) in Section 6.2 below.
(5-39) Soqotri (Semitic; Yemen)
a. \[
  [(le \quad ja\text{-}se\text{-}l\text{-}zen \quad birhe)]_{\text{NP}} \text{nmlz}\_\text{pl.m.imp-love children come.3pl.m.per}
\]
'Ones (masc) who love children came.'
b. \[
  \text{men NMLZR.pl pl.m.imp-love children come.3pl.m.per}
\]
'Men who love children came.'

(5-40) German
a. \[
  \text{Ich treffe [den, [der morgen kommt]}_{\text{NP}} \text{nmlz}\_\text{sub.nmlzr tomorrow comes}
\]
'I meet the one who comes tomorrow.'
b. \[
  \text{Ich treffe [den Mann, [der morgen kommt]}_{\text{NP}} \text{nmlz}\_\text{man sub.nmlzr morning comes}
\]
'I meet the man who comes tomorrow.'

(5-41) Spanish
a. \[
  \text{[El [que está leyendo un libro]}_{\text{NP}} \text{nmlz}\_\text{es mi padre.}}
\]
'Vee the one who is reading a book is my father.'
b. \[
  \text{[El hombre [que está leyendo un libro]}_{\text{NP}} \text{nmlz}\_\text{es mi padre.}}
\]
'The man that is reading a book is my father.'

(5-42) Kanuri (Nilo-Sahalan; Hutchison 1981)
a. \[
  [\text{àwó}^{27} \text{[nyìà galàngîn]} = dü]_{\text{NP}} \text{nmlz}\_\text{fàné!}
\]
'Thing to,you 1sg.advise=det 2sg.listen.IMV
Listen to what I am advising you.'
b. \[
  [kâm \text{[rűk à nà= dü]}_{\text{NP}} \text{sàwàːnàm}
\]
'Person 1sg.saw=det your.friend
The person that I saw is your friend.'

(5-43) Akan (Kwa; Campbell 2013)
a. \[
  \text{Mè\text{-}n\text{-}hù [dèè [ษ-bì-ษ màmè] nò]}
\]
'1sg.neg-like NMLZR 3sg.pst-hit woman def
I didn’t like the one who hit the woman.'
b. \[
  \text{Pàpà nò [dèè [mè\text{-}hyiè nò ènìlà] nò]}
\]
'man def NMLZR 1sg.meet him yesterday def
'the man that I met yesterday'

\[27\] See Section 8 on this marker.
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(5-44) Chichewa (Bantu; Mchombo 2004, p. c.)

   10-NMLZR 2.PL-PRS-10OM-want-fv 10-be 16.LOC 7.anthill
   ‘What you want is on the anthill.’

   10.goats 10-NMLZR 2.PL-PRES-10OM-want-fv 10-be 16.LOC 7.anthill
   ‘The goats that you want are on the anthill.’

Many more similar examples could easily be adduced, but the above is perhaps enough to dispel the widely-accepted analyses of so-called relative clauses as independent structures apart from nominalizations, analyses that give rise to limited observations such as “a somewhat more rare function of nominalization: as a relative clause modifying a head noun” (Comrie & Thompson 2007: 378) or claims like “in certain languages relativization is indistinct from nominalization” (Comrie & Thompson 2007: 379) that suggest that relativization and nominalization are in principle two distinct structures.28

It is worth noting that the NP-use of argument nominalization plays an important role in the formation of Wh-questions and so-called cleft (or focusing) constructions in a fair number of languages. Perhaps the most well-known of this is the case in Western Malayo-Polynesian, where we find the following patterns (see Bruil (this volume) for a pertinent discussion):

(5-45) Sasak (Pancor dialect, Western Malayo-Polynesian)

Argument nominalization in NP-use

a. [Si Ø mbace buku=ne] batur=ku.
   NMLZR AF.read book=this friend=1SG
   ‘The one who read this book is my friend.’

Argument nominalization in modification-use

b. Dengan [si Ø mbace buku=ne] batur=ku.
   man NMLZR AF.read book=this friend=1SG
   ‘The man who read this book is my friend.’

Wh-question

c. Sai [si Ø mbace buku=ne]
   who NMLZR A.read book=this
   ‘Who read this book?’/(lit.) ‘Who is the one who read this book?’

28. Views much wider than these and that are consistent with our analysis have been expressed by those working on Tibeto-Burman languages. DeLancey (2002: 56), for example, notes that “[t]he fundamental relativization pattern is the same throughout the family: relativization is a subspecies of clausal nominalization. The modifying clause is nominalized, and then stands in either a genitive or appositive relation to the head noun.” Compare Newar examples (3-16) and (7-19) in the text illustrating DeLancey’s point. See also Noonan (1997) and (2008).
Cleft
d.  \(\text{Loq Ali [si } \emptyset \text{ mbace buku=ne]}\)
   \(\text{ART Ali NMLZR A.read book=this}\)
   ‘It is Mr. Ali who read this book./ (lit.) ‘Mr. Ali is the one who read this book.’

(5-46) Yaqui (Southern Uto-Aztecan; Alvarez 2012: 89, p.c.)
a.  Wh-question
   \(\text{Jabesa [wa } \emptyset \text{ jiosam noktua-me]}\)
   who \(\text{DEM book read-NMLZR}\)
   ‘Who is reading the book?/ (lit.) ‘Who is the one that is reading the book?’

b.  Cleft
   \(\text{Joan [wa-me } \emptyset \text{ yabe-m tea-ka-me]}\)
   \(\text{John DEM-PL key-PL find-PERF-NMLZR}\)
   ‘It is John who found keys.’ (lit.) ‘John is the the one who found those keys.’

(5-47) Rainy River Ojibwa (Algonquian; Johns 1982)
a.  Wh-question
   \(\text{wenen [}[\text{ka}:\text{bimpato:t}]_{\text{NMLZ}}\text{]}_{\text{NP}}\)
   who \(\text{NMLZR (pst).run.3c}\)
   ‘Who ran?’/(lit.) ‘Who is the one that ran?’

b.  Cleft
   \(\text{ikwe [}[\text{ka}:\text{nagam}\text{at}]_{\text{NMLZ}}\text{]}_{\text{NP}}\)
   \(\text{woman NMLZR.sing.3C}\)
   ‘It’s the woman who is singing.’/ (lit.) ‘The woman is the one who is singing.’

(5-48) Thompson River Salish (Kroeber 1977: 387)
a.  Wh-question
   \(\text{Swet [k } [\text{wik-t-x}]_{\text{NMLZ}}\text{]}\)
   who \(\text{ART see-TRZ-2S.TS}\)
   ‘Who did you see?’/ (lit.) ‘Who is the one that you saw?’

b.  Cleft
   \(\text{c’ew’stin [e } [\text{n-s-tx}^\text{a}]_{\text{NMLZ}}\text{]}\)
   \(\text{soap ART 1SG.POSS-NMLZR-BUY}\)
   ‘It is soap that I bought./ (lit.) ‘Soap is what I bought.’

Before turning to the theoretical discussions of these usage patterns of grammatical nominalizations, let us take a quick look at another modification-use of grammatical nominalizations, namely their adverbial use. The adverbial use of grammatical nominalizations typically occur together with a conjunctive particle.
or postposition, as in the Mekens and Mëbengokre examples below (see also van Gijn, this volume), but a straightforward use of nominalizations in adverbial function is also seen in the Americas, as illustrated by some examples below, and elsewhere (e.g. Japanese, Sasak, Gã). As the label for each example below shows, these adverbials denote/refer to such concepts as purpose, condition, time, place, and manner metonymically evoked by the events that the relevant event nominalizations denote (see Van Linden, this volume, for a range of adverbial function of nominalizations in a single language).29

(5-49) Tarma Quechua (Quechuan; Adelaar 2011)

Purposive:  
[yaku picha-q]-mi away-ya-
water clean-NMLZR-AF go-PR-1SG  
'I am on my way to clean the water (canal).'

Conditional: [mana nuqmchik kuga-ta traqtra-sha]-m kiru-nchi ismu-n  
not we [INCL] coca-ACC chew-NMLZR-AF tooth-4P rot-3SG  
'Our teeth rot if we do not chew coca.'

(5-50) Tapiete (Tupí-Guaraní; Ciccone 2008)

Purposive:  
[1-1.4] she a-ha a-heka [a-mbr’invita-wërâ]  
I 1.SG.AC-go 1.SG.AC-look.for 1.SG.LOAN-invite-NMLZR  
'I go to look for (fish), in order to invite (you).'

(5-51) Kakataibo (Panoan; Zariquiey 2011)

Time:  
xu ‘ikë [xu ‘i-kë] kana ’êx Lima-nu kwanakën  
small be-NMLZR.NAR.1SG 1SG=L Lima-LOC go-rem.pst-1/2  
'I went to Lima when I was small.'

(5-52) Mekens (Tupí; Galucio 2011)

Temporal/conditional:  
kiri se-ayt-kwa-t [se-akar-ab]=ese  
child 3C-cry-TR-PST 3C-fall-NMLZ=LOC  
'The child cried when he fell down.'
(lit.) 'The child cried at his own falling.'
Cause (reason)/time:  
ôt o-akara [ôt o-etayap-ka-ab]=ese  
I 1SG-fall I 1SG-slip-TR-NMLZ=LOC  
'I fell down because I slipped; I fell down when I slipped.'

29. Recall that these concepts are also denoted by nominalizations that function as the head of NP-arguments (see (3–14)). Nominalizations in adverbial use are adjuncts ([I will try to be home [when you arrive]]) rather than arguments ([I want to know [when you arrive]]).
Simultaneous event:

\[
\begin{align*}
[a-\text{je} & \text{ tep } b\text{ör}]\quad n\text{yr}y\text{r}i\ d\text{ja} & ba & ngôj ku\text{’}ô \\
2-\text{ERG fish} & \text{ roast.PL.NMLZR while} & \text{ FUT I.NOM pot} & \text{ wash}
\end{align*}
\]

‘I’ll wash the pots while you roast the fish.’

Simultaneous event:

\[
\begin{align*}
N\text{üü} & [k\text{ahni } t\text{ünto’è-tü}]_{\text{NMLZ}} p\text{ahekkawa}. \\
\text{I} & \text{ house climb-NMLZR} & \text{ fell}
\end{align*}
\]

‘Climbing on the house, I fell off.’

Reason:

\[
\begin{align*}
p\text{uyt} & \text{ ku } t\text{ e} & [n-s-q\text{àaz}] \\
\text{lie} & \text{ 1S.IS OBL ART} & \text{ 1S.POSS.NMLZR-tired}
\end{align*}
\]

‘I lay down because I got tired.’

Closely related to these adverbial uses of nominalizations is their use in so-called clause-chaining, serial verb, or converb constructions, which are typically translated into English conjoined sentences connected by \textit{and}, as adverbial forms in sequence as \textit{after doing X, after doing Y} ... or as participial forms in sequence as \textit{having done X, having done Y} ..., as seen in the translation of the following Japanese example.

(5-55) Thompson River Salish (Kroeber 1997: 381)

Reason:

\[
\begin{align*}
p\text{uyt} & \text{ ku } t\text{ e} & [n-s-q\text{àaz}] \\
\text{lie} & \text{ 1S.IS OBL ART} & \text{ 1S.POSS.NMLZR-tired}
\end{align*}
\]

‘I lay down because I got tired.’

The \textit{-i/-Ø} nominalizations here are the grammatical counterparts of the stem nominalizations discussed in Section 2.2, and the structures marked by them, like the chained structures in other languages, lack tense and other finite verbal marking, unlike the tense-marked final verb \textit{si-ta ‘do-pst’}.

While the chained structures involving stem nominalizations seen in (5-56) is occasionally heard, they are mainly found in written form. As below, the modern colloquial chain structure in Japanese involves additional marking by so-called conjunctive particle \textit{-te}, whose use in chain structures goes back to Old Japanese (8\textsuperscript{th} C).
(5-57)  Taroo wa  [tosyokan ni it = te (< ik-i = te)], [hon o
taro top library to go-nmlzr = con book acc
yon = de (< yom-i = te)], [siryoo o sirabe = te (< sirabe-O = te)],
read-nmlzr = con material acc check = con
syukudai o si-ta.30
homework acc do-pst
‘Taro, having gone to the library, having read the book, (and) having checked
out the material, did the homework.’

Similar uses of nominalizations as chains are widely observed among the languages
of the Americas. For example, Northern Paiute has a nominalizing suffix -na, which
(a) yields action nominals from verbs, (b) marks object nominalizations in argument
nominalization, and (c) which marks event nominalizations functioning as verb
complements; (d) it also marks structures chained in the manner of the Japanese
examples above. Thornes (2003: 128–129, 443) nicely illustrates these diverse uses
of na-nominalizations in Northern Paiute as below, where his original gloss for -na
in grammatical nominalizations follows the traditional term “participle” (PTCP).

(5-58)  Northern Paiute (Numic, Uto-Aztecan)
a.  Action nominal
i=nossi-na waha-na
= dream-nmlzr tell.of-nmlzr
‘…telling of my dreams …’ (telling my dreaming)
b.  Object argument nominalization (in modification-use)
su=miidi  [i=kuhani-na] kai toki kamma
nom=meat 1=cook-nmlzr neg correct taste
‘The meat I cooked doesn’t taste right.’ (lit. “the meat of my cooking”)
c.  Object complement
u-su  3-nom  ka=nimizdoho u=nagi-kya-na] puni
3-nom obl=people,Masher 3=chase-trnsl-nmlzr see
‘S/he saw the Nemedzoho chasing her/him.’
d.  Chained nominalization
yaisi isu kaiba kussi timatai-na yaisi usu pabiẓi
then this mountain dust rise-nmlzr then that elder.brother
u-punni-na uka kussi-ba yaisi pisa u=supidakwatu
3=see-nmlzr that.obl dust.loc? then well 3=understand
‘…then as the dust rose from this mountain, and the elder brother
(Wolf) saw it, that dust, then (he, Wolf) understood it (i.e. that it
meant).’

30. The parentheses enclose the historically antecedent forms, from which the modern forms
arose via phonological changes.
Languages that allow nominalizations without specific nominalizing morphology may use unmarked nominalizations in their chaining constructions. Creek has several nominalizing suffixes, some of which can be used in marking both lexical and grammatical nominalizations. In addition, the language allows grammatical nominalizations that are not specifically marked as such. For example, Martin (2011) analyzes the bracketed structures below as involving no nominalizing morphology.31

(5-59) Creek (Martin 2011: 392, 394)

a. hatâm [pɔ̆ːmi ɔː:i-ːn] pɔ̆ːn-homp-ick-in
   again we have.FGR-1PL.AG-REF.N 1PL.DAT-eat.LGR-2S.AG-N
   om-áti-s be-FUT-IND
   ‘Then you will eat [what we have]…’

b. asêy ifá [aː-höyt-aːti-t] lopéye-iːt ɔː:s
   that dog DIR-stand.FGR-REF-T nice-DUR-T be.FOR-IND
   ‘That dog [standing over there] is friendly.’

We know that the structures in the brackets above have been or are nominalized or nominal because they are marked by the referential markers -a:n and -a:ti, which mark referential or contrastive nominals, where the former combines with the object case marker -(i)n and the latter is followed by the subject case marker -(i)t. It is this type of case-marked nominal structures that form chains in Creek, as seen below, where the case markers take on a switch-reference function, with the subject marker -(i)t marking same subject across chained structures and the object marker -(i)n different subject.32

(5-60) Creek (Martin 2011: 346)

aː-oséyy-in, halâ:t-ey-n, an-ciyałl-in,
dir-come.out-HGR-N hold-FGR-1S.AG-N 1S-DAT-struggle.against-LGR-N
 tí-weyk-éy-n, hola:n-ít
rcp-throw.LGR-1S.AG-N defecate.LGR-T

‘[Rabbit] came out [DS], I grabbed him [DS], he struggled against me [DS], I threw him down [DS], and he crapped [SS]…’

31. There is a possibility (if historical) that referential markers a:n and a:ti seen below are connected with the agentive nominalizer -a in the language.

32. The development of switch-reference markers out of nominalizers appears to be a widespread phenomenon. See Jones & Jones (1991: Section 11.2) on Barasano and the contributions to this volume by Zariquiey and Valle & Zariquiey on Kakataibo.
Contrary to the widely used term “clause-chaining”, the chained structures are not predicing clauses. They are in fact event nominalizations denoting events that hold temporal relations (sequential or simultaneous) among themselves or with regard to the asserted event.33 Chained nominalizations may have an overt marker indicating the temporal relations or may not. In the case of Japanese (see (5-56)–(5-57)) and Creek, the order of juxtaposition of nominalizations indicates a sequential temporal relation, whereas in Northern Paiute (5-58), nominalized structures in chain by themselves denote simultaneously occurring events. Japanese uses the marker -nagara for simultaneous events. Northern Paiute, on the other hand, marks sequential events by means of the suffix -si, as below, which contrast with the simultaneous events seen in (5-58d).

(5-61) Japanese
\[
\text{Taroo wa [aruk-i-nagara] hon o yonda.}^{34} \\
\text{Taro TOP walk-NMLZR-SIMUL book ACC read} \\
\text{‘Taro read while walking.’}
\]

(5-62) Northern Paiute (Thornes 2003: 457)
\[
\text{mi=tika-k i-u-si nimmi tiwau mia-si na-noo-ka-si oo} \\
\text{use=eat-APL-PMC-SEQ WE.EXCL again go-SEQ MM-CARRY-TRNSL-SEQ DEM} \\
\text{ka=yamoso-tami} \\
\text{obl=Ft.Bidwell-toward} \\
\text{‘Having allowed us to eat, we went on again, and were hauled on out there to Fort Bidwell …’}
\]

We have presented above the general usage patterns of grammatical nominalizations with a special focus on the two uses of argument nominalizations across a wide variety of languages of both South and North America, as well as elsewhere across the world. Having laid out the empirical foundations of the usage patterns of grammatical nominalizations, we are now in a position to launch theoretical discussions. We start with the NP-use of event nominalizations.

5.2 NP-use of event nominalizations: So-called “internally-headed RCs”

One of the major issues pertaining to the use of event nominalizations centers around the construction in (3-3d). Comparison of a similar example and an ordinary relative clause construction in Quechua below illustrates the problem at issue.

33. See Section 6.2 on the definitions of clauses, sentences, and nominalizations.

34. Unlike sequential events like (5-56) and (5-57), the -nagara form allows only a single simultaneous event per sentence.
(5-63) Bolivian Quechua
a. \([Maria \text{wallpa-ta} \ wayk'u-sqa-n]-ta \ mik'u-sayku\)
   Maria chicken-ACC cook-NMLZR-3 ACC eat-PROG.1PL.EXCL
   ‘We are eating the chicken that Maria cooked.’
   (lit.) ‘We are eating Maria cooking the chicken.’

b. \([\text{Maria} \ O \ wayk'u-sqa-n] \text{wallpa}-ta \ mik'u-sayku\)
   Maria cook-NMLZR-3 chicken-ACC eat-PROG.1PL.EXCL
   ‘We are eating the chicken that Maria cooked.’

Following Gorbet’s (1974) lead, a large number of researchers (Keenan 1985, Cole 1987, Kuroda 1992, etc.) have analyzed the constructions similar to (5-63a) as “internally-headed relative clauses”, assuming (i) that they are relative clauses and (ii) that a head nominal exists within “relative clauses” unlike ordinary relative clause constructions, where a head exists externally in the main clause. A problem with the first assumption is that it is not at all obvious that these structures have the function of relative clauses, which is either to restrict the denotation of the head noun (restrictive relatives) or to identify the head noun in terms of the denotation of a modifying nominalization structure (non-restrictive relatives). That the idiomatic English translations of the relevant structures turn out to be relative clauses is hardly acceptable evidence for the proposed internally-headed RC analysis. Indeed, my Quechua consultant from the Cochabamba village in Bolivia would use the externally-headed RC in (5-63b) over the so-called internally-headed RC in (5-63a) in answering a question such as “What/which chicken are you eating?”

Actually the only evidence that suggests the relative clause status of the so-called internally headed RCs is the fact that similar, but not identical, structures are used for externally-headed relative clauses. There are, however, languages that are said to have internally-headed RCs in the absence of externally-headed RCs (e.g. Yuman languages Diegueño, Jamul Tiipay, and perhaps others, Kutenai, Seri, Parkatêjê), showing a measure of independence of the two.

The second assumption that in these structures an argument internal to the “relative clause” is the argument of the main-clause predicate is also problematic. Those who have studied so-called internally-headed RCs have not looked at constructions like (3-3e), where there is no NP within the “relative clause” that can serve as a main-clause (semantic) argument; accordingly they cannot be analyzed as internally-headed RCs. Notice the exact structural parallelism between so-called internally-headed RCs and the resultative nominalizations below:

(5-64) Bolivian Quechua
a. \([Maria \text{wallpa-ta} \ wayk'u-sqa-n]-ta \ mik'u-sayku\)
   Maria chicken-ACC cook-NMLZR-3SG ACC eat-PROG.1PL.EXCL
   (lit.) ‘We are eating Maria cooking the chicken.’ ‘We are eating [the chicken involved in the event of] Maria’s cooking a chicken.’
b. [Maria laranjas-ta ch’irwa-sqa-n]-ta ujya-ni
   Maria oranges-ACC squeeze-NMLZR-3SG-ACC drink-1SG
   (lit.) ‘I drink that Maria squeezed oranges.’
   ‘I drink [the result of] Maria’s squeezing oranges.’
   Cf. *laranjas-ta ujya-ni
   oranges-ACC drink.1SG
   ‘I drink oranges.’

In (5-64b) it is clear that the nominalization-internal argument (laranjast ‘oranges’) is not the semantic argument of the main-clause verb; yet the sentence is perfectly acceptable. Analyzing the form in (5-64a) as an internally-headed relative clause leaves resultative nominalizations of the type seen in (5-64b) unaccounted for.

Those who subscribe to the internally-headed RC analysis have not bothered to expand their data beyond the earlier observations and thus fail to recognize the likely fact that those languages permitting so-called internally-headed RCs allow resultative nominalizations of the type seen above. Besides Bolivian Quechua, in three more languages allowing so-called internally-headed relative clauses for which native speakers were available to the present author, this prediction turned out to be correct, as evident from the following examples from Japanese, Northern Qiang (Tibeto-Burman; China), and Parkatêjê in northern Brazil.

(5-65) Japanese

Ken wa [Hana ga mikan o sibotte kureta] no o hitoiki
Ken TOP Hana NOM orange ACC squeeze.GER gave NPM ACC one.gulp
ni nonda.
in drank
(lit.) ‘Ken drank that Hana squeezed oranges for him in one gulp./Ken drank [the resultant product of] Hana’s squeezing oranges for him in one gulp.’
Cf. *Boku wa mikan o nona.
   I TOP orange ACC drank
   ‘I drank oranges.’

(5-66) Northern Qiang (Tibeto-Burman; courtesy of Chenglong Huang)

a. [themle-wu tcysståsi ha-tśa-tha-ji] lo-qu qa
   3PL-AGT orange-juice DIR-squeeze-PART-CSM DEF-CLF 1SG
   sso-tsha.
   DIR-drink.1SG (CLF = classifier)
   ‘I drank [the orange juice involved in] their squeezing orange juice.’

   3PL-AGT orange DIR-squeeze-PART-CSM DEF-CLF 1SG DIR-drink.1SG
   ‘I drank [the resultant product of] their squeezing oranges.’
Our claim is that, in both so-called internally-headed RCs and resultative nominalizations of the above type, the semantic arguments of the relevant predicates are not to be sought structure internally in a direct manner, as is done by the proponents of the internally-headed RC analysis, because they are actually something that are evoked metonymically. In fact, many event nominalizations in NP-use lack internal arguments functioning as semantic arguments. For example, the semantic argument of the main-clause predicate in [John's falling off of the bed] happened at 3:00 am is the event (of John's falling off of the bed) evoked by the nominalization structure, but the noun event is not found anywhere in the relevant structure. Similarly, the semantic object argument of the verb know in I know [John is honest] is a fact (that John is honest), not the state of affairs of John's being honest, but again there is no noun fact found in the structure; cf. the synonymy between I know John is honest and I know the fact that John is honest.

What we are seeing here are cases of the metonymy-mediated syntax-semantics mismatch, as seen in ordinary metonymic expressions of the following type: *The first violin is sick today, I heard three CDs tonight*. The metonymy-based nominalization analysis proposed in this paper treats all these constructions as nominalizations that, like lexical nominalizations discussed earlier, metonymically evoke concepts such as facts, propositions associated with events or state of affairs at large, event protagonists/participants, resultant products, as well as circumstantial matters like time, location, and reason closely associated with an event. In these constructions, nominalization structures function as syntactic arguments as a subject or object precisely because they evoke and stand for thing-like entities just like ordinary nouns do. And it is these metonymically evoked entities that function as semantic arguments of the main predicates. We are arguing that all the structures in (5-68) and (5-69) below display the same syntax-semantic mismatch mediated by metonymy as indicated below, where {...} represents what is evoked metonymically and what functions as semantic arguments.

35. Thanks are due to Marília Ferreira for checking the Parkatêjê forms with native speakers.
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(5-68)  

a. The first violin [PLAYER OF] is sick today.

b. I heard three CDs [MUSICAL SOUNDS OF] tonight.

(5-69)  

a. FACT

English: Bill knows [that John is honest]_{nmlz} \{EVENT > FACT\}

b. EVENT PROTAGONISTS (so-called internally-headed relative clauses)³⁶

Slave

[li gah hedēhfe] i [EVENT > PROTAGONISTS] ghāyeyidā
dog rabbit 3.chased nmlzr

i. ‘I saw the dog that chased the rabbit.’

ii. ‘I saw the rabbit that the dog chased.’

c. RESULTANT PRODUCT (resultative nominalization)

Waiwai

[a-mok-ri] [EVENT > RESULT (sounds)] w-enta
2-come-AC.nmlzr I-hear+IMM.pst

‘I heard you/your coming.’

d. PLACE

Mosetén (Sakel 2004: 95)

chhiko’-ni-ti-dye’ [EVENT > PLACE]
liquid-put-VD-nmlzr
‘places where one washes oneself’

Metonymic meaning extension often works transitively such that one metonymically evoked concept leads to another closely related concept (cf. the red, white, and blue > [THE U.S. NATIONAL FLAG] > [U.S.A.] in Team USA players …representing the red, white and blue at the 2014 FIL World Championship …). It is assumed that event nominalizations of the kind seen above are similar in that the nominalization structures first evoke events portrayed by the relevant structures, which in turn evoke those concepts associated with events, as indicated above. Notice how so-called internally-headed RCs and the resultative nominalization receive a uniform treatment under our analysis, while the internally-headed RC analysis leaves the latter unaccounted for.

³⁶. Our account does not explain why certain languages more readily allow this type of nominalization than others, where the so-called internally headed RCs do not obtain.
As shown by such forms as *I know that John is honest*, *I heard John sing in the shower* and *I saw the man kiss my daughter*, English allows syntax-semantics mismatches too. What is remarkable about those languages permitting so-called internally-headed RCs is the extent to which similar mismatches are allowed, permitting the equivalents to *I recorded John sing in the shower*, *I drank Mary squeeze the oranges* and *I scolded the man kiss my daughter* and *The man kiss my daughter is a flirt.*

5.3 NP-use of event nominalizations: So-called “complement clauses”

A subject complement (5-4a) and an object complement (5-4b) are also known as verb complements, as opposed to noun complements in (5-5). The traditional name for noun complements is “content clause”. These traditional descriptions are generally maintained by contemporary researchers as can be seen from Dixon’s (2006) and Noonan’s (2007) definitions below.

In many languages certain verbs—notably ‘see’, ‘hear’, ‘know’, ‘believe’, ‘like’, and often also ‘tell’ and ‘want’—can take a clause, instead of an NP (noun phrase), as a core argument. This is called a complement clause. (Dixon 2006: 1)

By complementation, we mean the syntactic situation that arises when a notional sentence or predication is an argument of a predicate. For our purposes, a predication can be viewed as an argument of a predicate if it functions as the subject or object of that predicate. (Noonan 2007: 52)

As indicated by his remark “[l]anguages lacking a full range of complement clause constructions will often employ some kind of nominalization as a complement strategy”, Dixon (2006: 37) recognizes a use of nominalizations as complements. The following show that languages of the Americas do indeed use nominalizations as verb complements.
(5-71) Tapiete (Ciccone 2008)

\[a\text{-}m\text{-}b\text{-}\text{e}u\text{-p}\text{o} \quad [y\text{a}w\text{a} \ y\text{i}\text{-w}\text{a}r\text{e} \ h\text{a}n\text{d}\text{i} \ a\wá\text{r}a\text{-}w\text{a}]\]
1.sg.ac-tell-FUT tiger 3-play with fox-NMLZ
'I want to tell that the tiger played with the fox.'
or 'I want to tell about the tiger that played with the fox.'

(5-72) Barasano (Tucanoan; Jones & Jones (1991: 160))

\[w\text{u}\text{-r}\text{-}k\text{a} \ t\text{i} \ e\text{h}\text{a}\text{-r}o\text{-t}i\text{-r}\text{e} \ b\text{a}\text{s}\text{-}i\text{-b}\text{e}\text{-a}\text{-h}\text{a} \ y\text{u}\]
fly-PTCPL-hollow 3in arrive-FUT-NMLZ~PROX-O know-NEG-PRT~3 1SG
'I don’t know when the plane is to arrive.' Or ‘I don’t know whether the plane will arrive or not.’

(5-73) Tariana (Arawakan; Aikhenvald 2009: 201)

\[w\text{a}\text{-}d\text{a}\text{l}\text{i}p\text{a} \ p\text{h}\text{i}j\text{i} \ d\text{i}\text{-}n\text{u}\text{-}r\text{i} \ p\text{h}\text{e}\text{m}\text{a}\text{-k}\text{a}\text{-n}\text{a}k\text{a}\]
1PL-towards agouti 3SGNF-COME-NMLZ IMP.HEAR-DEC-PRT.VIS
‘One can hear an agouti come towards us.’

(5-74) Shoshoni (Northern Uto-Aztecan; Dayley 1989: 274)

\[N\text{üü} \ [k\text{u}n\text{a}\text{i} \ w\text{a}y\text{a}n\text{t}\text{ü}n\text{n}n\text{a}] \ p\text{u}n\text{k}k\text{k}a.\]
I wood.O burn.NMLZ.O see
‘I see the wood burning.’

(5-75) Thompson River Salish (Kroeber 1977: 381)

\[Nm\text{u}\text{t} \ x\text{‘o}x\text{’} s\text{-}m\text{‘}n\text{-x}\text{‘} \ [k \ s\text{-}c\text{‘}q\text{’}\text{-}t\text{-}e\text{x}\text{‘}] \ n\text{m} \ w\text{a}nt\text{-}r\text{l}\text{t}\text{-}2\text{~}t\text{s} \ a\text{r}t \ NMLZ\text{-w}r\text{i}t\text{e}\text{-}T\text{r}\text{z}\text{-}2\text{~}t\text{s}\]
‘(and) you want to write it.’

A major issue here is whether what is known as clausal/sentential complements are structures that need to be recognized as distinct from what we have identified above as grammatical event nominalizations. Dixon’s (2006) comment quoted above makes it clear that he considers verb complements to be clauses and believes that they are different from nominalizations. The real question here is whether the formal differences between “complement clauses”, defined by Dixon as those showing a high degree of structural parallelism with clauses/sentences, and nominalizations make any substantial difference beyond the structural differences.37

37. See also the discussions by van Dijk, Haude & Muysken (2011: 3) in their introduction to the recent volume on subordination in South American languages, where these editors endorse Dixon’s (2006) division between nominalizations and subordinate clauses by saying that “[n]ominalization in particular is a common strategy that South-American languages use in the same places where other languages have subordinate clauses.” Notice the caution taken by Noonan (1985) in the earlier quote, where he says that complements obtain “when a notional sentence or predication is an argument of a predicate” rather than asserting that complements are clauses or sentences. (Emphasis added).
This is similar to asking whether the formal difference between the two lexical nominalizations \(a\) speaker and \(a\) cook makes any significant difference beyond the formal difference. If what Dixon considers to be complement clauses and nominalizations were two fundamentally different linguistic units, one would have to ask why they function in a similar way, both denoting events, facts and other abstract substantive, and both standing in a subject and object position that are typically occupied by nouns.\(^{38}\) This is the crux of the issue that needs to be addressed. Our position is that all those that function as verb and noun complements are nominalizations, whether or not they contain discernible nominalizing morphology or they show structural resemblances to clauses/sentences.

A deeper issue lying beneath the distinction between “complement clauses” and “nominalizations” Dixon draws depends on one’s understanding of what nominalization is and upon his analysis. Dixon’s understanding of what nominalization is is stated as below:

> ‘Nominalization’ is used to describe a process (and its result) by which something with the properties of a nominal can be derived from a verb or adjective, or from a complete clause. (Dixon 2006: 36; emphasis added)

Unfortunately this understanding does not help us distinguish between what Dixon calls “complement clause” and a nominalization since the latter, according to him, may also have structural properties of a complete clause; and they indeed do (see (5-71) and (5-72) above) like what Dixon considers to be a “complement clause”, e.g. John’s playing the national anthem (pleased Mary). One wonders why this example is not a (grammatical) nominalization because the structure certainly has “the properties of a nominal” in denoting an activity (see Dixon 2006: 15) and in syntactically functioning as a subject like an ordinary noun.\(^{39}\) The two examples Dixon gives in the section on nominalization strategy give the impression that he considers nominalizations those that have a nominalizing morphology, but then one wonders why the -ing suffix seen in the English “complement clause” example above is not nominalization morphology.

Dixon would consider the English that-construction in the translation of the Jamul Tiipay example below as a complement clause rather than a nominalization,

---

\(^{38}\) The same question must be posed for those who think that relative clauses are clauses distinguishing themselves from nominalizations (Comrie & Thompson 1985/2007); i.e. why do two fundamentally different structures function alike as noun modifiers if they were?

\(^{39}\) Dixon (2006: 15) discusses distinctions between this example and the expression John’s playing of the national anthem in terms of the distinction between a complement clause and a nominalization, but the pertinent distinction is between a grammatical nominalization and a lexical nominalization.
perhaps following the Generative Grammar tradition that labels such *that* as a complementizer (COM), a practice followed by many other contemporary grammarians, whereas the original Jamul Tiipay form would likely be considered a use of nominalization (“a nominalization strategy”) because it has the nominalizing morphology *-ch*, which points to a realis state of affairs.

(5-76) Jamul Tiipay (Yuman, Miller 2001: 223)

\[
\begin{align*}
\text{[\text{mi-iima-ch}_\text{NMLZ}-\text{pu}]_{\text{NP}} uuyaaaw} \\
2\text{-dance-NMLZR-DEM know}
\end{align*}
\]

‘He knows that you dance.’

However, the early generative grammarian Robert B. Lees, following the more traditional analysis, analyzes *that*-construction in the English translation above as a nominalization, “a factive nominal”, which he characterizes as an “abstract object” denoting “an abstract fact, or statement, or [as] information” (Lees 1963: 59ff). And to our mind, Lees is absolutely correct.40

Wisely, Miller (2001: Ch.7.2) treats all similar constructions as involving nominalization regardless of whether there is a nominalizing morphology, as in (5-75) above, or not, as in the example below.

(5-77) Jamul Tiipay (Miller 2001: 219)

\[
\begin{align*}
\text{[\text{me-xap}_\text{NMLZ}-\text{pu}]_{\text{NP}} nyawach my-uuwiw} \\
2\text{-enter DEM we.sj ½-see.pl}
\end{align*}
\]

‘We saw you come in.’

While the fact that both forms above display an external property of being marked by a demonstrative is a good indication that they are nominals heading an NP, Miller’s understanding of nominalization is functional and does not dependent upon morphology.41

40. Our major complaint about Lees’ work on English nominalizations is that he considers the relevant nominalization structures to be clauses (see Section 6.2). Robert B. Lees was the first PhD in linguistics at MIT, where he trained with Morris Halle and Noam Chomsky. He became a leading generative grammarian in the 1960’s, producing a number of influential linguists at the University of Illinois, Urbana-Champaign, including Ronald Langacker.

41. Dixon (2010: 316) also wants to distinguish between relative clauses and nominalizations, as is clear from his remarks: “In some languages, a verbal affix marking a relative clause is homophonous with a nominalizer. This should not be taken to mean that a relative clause is a type of nominalization”. Why are certain relative clauses are homophonous with nominalizers if they are distinct constructions? Is it accidental?
5.4 NP-use of argument nominalizations: So-called “headless/free relative clauses”

Argument nominalizations in NP-use, as in (5-78) and (5-79) below, are generally known as “headless relative clauses” or “free-relatives” as if they are a kind of, or a derivative of, relative clauses. As in the case of so-called internally-headed relative clauses, it is not obvious that the relevant structures function as a restrictive or non-restrictive modifier. There is no head noun that sets a domain of denotation which is restricted to a subset or which is identified and commented on by the denotation of the relevant structure.

(5-78) English

You should marry [[who Ø110 loves you]_{NMLZ}^{110}]_{NP} not [[who Ø125 loves your money]_{NMLZ}^{125}]_{NP}

(5-79) Navajo (Southern Athabaskan; Willie 1989: 435)

[[Ø15 ’at’ééd yizts’όs-yéé]_{NMLZ}^{15}]_{NP} yalt’

girl 3sO: 3sS:kiss-NMLZR speaking

‘The one who kissed the girl is speaking.’

Actually, there has been little argument for considering these as a type of relative clauses. Most simply assume that to be the case because the same structures are also used as externally-headed relative clauses. This is like putting the cart before the horse to our mind. But there have been, in fact, some proposals to make so-called headless relative clauses align with externally-headed relative clauses. One is a deletion analysis that posits an external head noun, which is then deleted. This is the most popular analysis of headless relative clauses that has been applied to a wide range of languages (e.g. Weber (1989) on Quechua, Huang (2008) on Qiang, Treis (2008) on Kambaata). Another is to posit some kind of phonetically empty pronoun as the head, as proposed by Jelinek (1987).

Besides the fact that the same structure is typically used as ordinary relative clauses, a basic motivation for these analyses appear to be the semantic properties associated with the relevant structures, namely they denote entity concepts. Apparently it has never occurred to the proponents of these analyses that so-called relative clauses themselves (e.g. English forms [who loves you], [who loves your money] or Navajo structure [’at’ééd yizts’όs-yéé] ‘one who kissed the girl’ above) have entity denotations associated with them, in the manner described in Section 4, for example. This is largely because these structures had been thought to be clauses or sentences, which certainly do not denote thing-like concepts as ordinary nouns do (see Section 6.2 below). By positing a head noun or a pronoun, the proposed analyses capture the fact that the relevant structures are associated with entity denotations.
The problem with these proposals is that neither the deletion analysis nor the pro analysis is a complete description until they provide an account as to when the putative deletion of the posited head noun applies or how the pro head is actually distributed and interpreted. In offering a complete description, these analyses must refer to context, similar to our analysis. Our argument is that if one has to refer to the context anyway, let the context handle the whole thing. Our analysis maintains that so-called headless relative clauses are grammatical argument nominalizations and that they themselves, not a putative head, have a set of entity concepts as their denotations as described in Section 4. The context of use then determines the referent most appropriate per the Gricean Cooperative Principle. The deletion and pro analyses have been conceived because the proponents of these analyses lacked proper understandings of the nature of grammatical nominalizations and because of their usual practice of analyzing linguistic structures in complete isolation from the context in which they are used.

There are, furthermore, cases in which so-called headless relative clauses (our NP-use of nominalizations) obtain without possible externally-headed RC counterparts. In Russian nominalizations marked by the interrogative pronouns koto-royj ‘which’ and chto ‘what’ allow both NP-use and modification-use as shown in the examples below.

(5-80) Russian
a. tot, [kotoryj/cht0 stoit tam], eto drug otca
   that WHICH/WHAT stand there it friend father
   ‘The one who is standing there is my father’s friend.’

b. tot chelovek, [kotoryj/cht0 stoit tam], eto drug otca
   that man WHICH/WHAT stand there it friend father
   ‘That man who is standing there is my father’s friend.’

However, the nominalizations marked by kto ‘who’ allows only NP-use, as indicated by the ill-formed sentence in (5-81b) below.

(5-81) a. (tot,) [kto vymyl ruki], mozhet nachatj jestj
   (that) WHO42 washed hands can start eat
   ‘The one who has washed his hands can start eating.’

b. *Maljchik, [kto vymyl ruki], mozhet nachatj jestj
   boy WHO washed hands can start eating
   ‘The boy who has washed his hands can start eating.’

42. The glosses WHO, WHAT, etc. are employed to show the etymologies for these nominalizers.
In German whether or not nominalizations marked by *was* ‘what’ has a modification-use is dialectal. While High German is said to disallow a form like (5-82b) below, Bavarian German or some other southern German dialects may allow it.

(5-82) German

a. *Ich lese (das), [was Sie empfehlen]*
   
   I *read* ART WHAT you recommend
   ‘I’ll read what you recommend.’

b. *Ich lese das Buch, [was Sie empfehlen]*
   
   I *read* ART book WHAT you recommend
   ‘I’ll read the book which you recommend.’

On the other hand, in the case of those marked by *wer* ‘who’, a modification-use appears to be generally prohibited. Observe:

(5-83) a. *Ich empfange, [wer (auch) morgen kommt]*
   
   I receive WHO (also) tomorrow comes
   ‘I receive who(ever) comes in tomorrow.’

b. *Ich empfange den Mann, [wer morgen kommt]*
   
   I receive ART man WHO tomorrow comes
   ‘I receive the man who comes tomorrow.’

In Spanish *que*-marked nominalizations generally allow both NP-use and modification-use, as in (5-84) below, but those marked by *quien* lack a modification-use.

(5-84) Spanish

a. *Leeré lo [que usted recomienda]*
   
   I will *read* ART WHAT you recommend
   ‘I will read what you recommend.’

b. *Leeré el libro [que usted recomienda]*
   
   I will *read* ART book WHAT you recommend
   ‘I will read the book which you recommend.’

(5-85) a. *Veré a [quien viene mañana]*
   
   I will *meet* to WHO comes tomorrow
   ‘I will meet the one who comes tomorrow.’

b. *Veré al hombre [quien viene mañana]*
   
   I will *meet* to the man WHO comes tomorrow
   ‘I will meet the man who comes tomorrow.’

Thus, all in all, there is little motivation or evidence for analyzing the NP-use of argument nominalizations as (headless) relative clauses. The formal resemblances between so-called headless relative clauses and ordinary externally-headed relative is due to the fact that they represent two different uses of the same nominalization.
structures. The fact that so-called headless RCs denote and refer to entity concepts is not due to an external head that is doomed to be deleted or due to a phonetically empty pronoun; rather, it is because grammatical argument nominalizations underlying them themselves have such denotations allowing their referential use in actual discourse (see Section 4).

5.5 Modification-use of argument nominalizations: So-called “relative clauses”

As shown in (5-7), our analysis of relative clause constructions is straightforward. While it captures the observations on the use of nominalizations as relative clauses, it departs from the traditional analysis in several ways. A major claim advanced in this paper is that so-called relative clauses are all nominalizations, whether they contain finite verb forms or they contain specific nominalization morphology. As in the case of verb complements discussed in Section 3.1, those who believe that argument nominalizations and relative clauses are separate structures, which “in certain languages [are] … indistinct” (Comrie & Thompson 1985/2007: 379), must answer (i) why structurally they both have a gap (or a pronoun) in an argument or adjunct position (see Section 6.1), and (ii) why functionally they both modify a noun.

The traditional analysis of RC constructions based on English makes crucial reference to the role of so-called relative pronouns, such as who and which, that play the double role of indicating the dependency relation between the pronoun and a gap within an RC and of holding the perceived anaphoric relation with a head noun, giving rise to the term “relative pronoun”, as in the following representation.

(5-86) the man, [whom, [you love Øi]]

Such an analysis is problematic when applied to other languages in that many, if not most, languages do not use anything like relative pronouns.43 Most descriptions of RC constructions in a variety of languages label an element marking what looks like an RC as REL or as a relative pronoun. This practice, however, has not been independently justified in most of such descriptions; they simply follow the analysis of (5-86) based on English. To our mind, they are best analyzed as nominalizers, as indicated by our relabeling of them in the examples cited in this paper. Even in English, we can advance an argument for treating who, which, etc. as indefinite pronouns used as nominalizers or markers of nominalization (in addition

43. Because of this, some grammarians say that their languages do not have relative clauses (e.g. Jones & Jones 1991: 149). A more accurate way of saying this is that there are no English-style relative clause constructions in the relevant languages.
to their use as interrogative pronouns). The use of indefinite pronouns as nominalizers makes good sense because what argument nominalizations denote may be indefinite (e.g. [Who gets there first] gets the prize; You may choose [which you find most appealing]).

Our point is that relativization does not depend on so-called relative pronouns and that the perceived relation between so-called relative pronouns and gaps in RCs can be captured in terms of the role-marking morphology discussed in an earlier section. For example, the German nominalizers *der* and *den*, for example, mark the grammatical role of the entity denoted by an argument nominalization; they are a subject nominalizer and an object nominalizer that mark (or combine with) a subject nominalization and an object nominalization, respectively, as below.

(5-87) German

a. *Ich treffe [den Mann, [der [Ø morgen kommt]_{SUB,NMLZ}]_{NP}}
   ‘I meet the man who comes tomorrow.’

b. *Ich treffe den, [den [du mir vorgestellt hast]_{DO,NMLZ}}
   ‘I meet the to one whom you introduced to me.’

A requirement in languages with role-marking nominalization morphology like German is that the morphology correctly indicates the type of argument nominalization involved. That is, a subject nominalizer must combine with a subject argument nominalization with a gap in subject position, as in (5-87a), and an object nominalizer with an object argument nominalization with a gap in object position, as in (5-87b). The role of the English nominalizer *whom* is exactly the same as that of the German DO nominalizer *den* (except for the additional gender information coded in the latter). All languages with role-indicating morphology examined in Section 3.2 have similar requirements (see Section 9.1 for an important implication of these points in the analysis of relative clause constructions).

While many languages are similar to German in having role-indicating morphology, there are many others that do not; accordingly, such morphology, like so-called relative pronouns, is not an essential feature of RC constructions in general, as can be seen in the Japanese and the Toba pattern below.

(5-88) Japanese

a. *[Ø hon o yomu]_{SUB,NMLZ} kodomo]_{NP}
   book ACC read.prs child
   ‘a child who reads a book’

b. *[kodomo ga Ø yomu]_{OBJ,NMLZ} hon]_{NP}
   child NOM read.prs book
   ‘a book which a child reads’
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(5-89) Toba (Messineo & Porta 2009: 57)\(^{44}\)

a. so \([jiGaWa \{Ø i-waGan a-so qaʔapole\}]_{\text{SUB,NMLZ}}\)\(_{\text{NP}}\)
   \text{DD man 3A-hit FEM.DD young.lady}
   ‘the man who hit the young lady’

b. so \([jiGaWa \{a-so qaʔapole i-waGan Ø\}]_{\text{OBJ,NMLZ}}\)\(_{\text{NP}}\)
   \text{DD man FEM.DD young.lady 3A-hit}
   ‘the man whom the young lady hit’

The examples of RC constructions above indicate that a minimal requirement, the essential feature of RC constructions, is that they involve as a modifier an argument nominalization with a gap (or a pronoun as in Thai, Modern Hebrew and some others) in an argument (or an adjunct) position. The relevant argument nominalizations may or may not involve morphology indicating the grammatical role of the entity denoted by the nominalization. These considerations suggest the following analysis of RC constructions.

(5-90) a. Japanese
   \([\{Ø hon o yomu\}]_{\text{SUB,NMLZ}}\)\(_{\text{NP}}\) kodomo
   \text{book acc read.prs child}
   ‘a child who reads a book’

b. Chinese
   \([\{Ø zài nàr diào liyú\}]_{\text{SUB,NMLZ}} = \text{de} háizi\)\(_{\text{NP}}\)
   \text{PROG there fish carp = NMLZR child}
   ‘a child who is fishing carp there’

c. German
   [der Mann \([\{Ø dich liebt\}]_{\text{SUB,NMLZ}}\)]\(_{\text{NP}}\)
   \text{the man SUB.NMLZR you love}
   ‘the man who loves you’

d. English
   [the man \([\text{whom you love} \Ø]\)]_{\text{OBJ,NMLZ}}\(_{\text{NP}}\)
   OBJ.NMLZR

The above analysis embodies the idea that nouns have a denotation index in the manner of [dog]\(^{1}\) that points to a set of concepts that they denote. Nominalizations as nominals share this property, as indicated above. In the case of argument

\(^{44}\) Toba has the non-role marking nominalization marker \text{mazi}, which can be used at the beginning of an argument nominalization.
nominalizations, these denotation indices bind a variable in the form of a zero or pronominal argument, thereby indicating the grammatical role that the entity denoted by the nominalization structure stands for. In (5-90a), (b), (c), (e) above, the nominalizations denote an entity that stands for the subject role. In (5-90d) the nominalization denotes an entity that stands for the object role.

Notice that in our analysis there is no role that so-called relative pronouns play with regard to the gap in the nominalization or with regard to the head noun. This is a desired consequence of our analysis, which analyzes a restrictive relative clause construction as involving two independent nominals, each with its own denotation set. Restricting the denotation of the head noun means specifying its subset by the denotation of the modifying argument nominalization. Thus, the only requirement for the modifying nominalization with respect to the head noun in an RC construction is that the former denotes entities that intersect with those denoted by the head noun, as in Figure 3 below. Our analysis is highly compatible with the treatment of a restrictive relative clause construction in Formal Semantics, which would define the denotation of such a construction as the intersection of two sets of entities; e.g. \{x| x is a man\} \cap \{x| you love x\} (“the intersection of the set of all x such that x is a man and the set of all x such that you love x”), where x’s are two independent variables.\textsuperscript{45}

\[ \text{Figure 3. Denotation of restrictive relative clause} \]

As is clear from the exposition above, so-called subject relative clause construction is simply a combination of a head noun and a subject argument nominalization with a gap (or rarely a pronoun) in subject position (5-90c), (e), and so-called

\textsuperscript{45}. The Formal Semantic analysis would have a difficult time in deriving the second set for the modification involving event nominalizations without a gap (see Section 3.1), which would not yield to an analysis calling for operator movement, as in the case of the generative analysis of wh-relatives in English.
object RC construction a combination of a noun head and an object argument nominalization with a gap or a pronoun in object position (5-90d). The relativization process per se involves no movement or deletion of a pronoun or a noun coreferential with the head noun. It simply brings together, or merges, a head noun and a grammatical argument nominalization that restricts the denotation of the head noun in the manner shown in Figure 3.

In addition to the distinction between the restrictive and the non-restrictive relative clause constructions, there are relevant constructions that deviate from the canonical RC constructions discussed above. First, many languages of the world do not seem to make a clear distinction between restrictive and non-restrictive RCs as in English, where the latter is set off by a brief pause in speech and a comma in writing. Nevertheless, there is a functional difference between the two even if the two constructions are formally alike. Instead of restricting the denotation of the head noun to its subset, the denotation of a modifying argument nominalization in non-restrictive RC constructions identifies that of the head noun either under strict identity, that the two denotations denote an identical entity, or as an instance of of the latter. For example, in the construction the man, who you have decided to marry against my advice, (is a real crook), the modifying argument nominalization likely denotes a single-member set. In such a case the total identity obtains between the denotation of a single-member set of the head noun and that of a single-set member of the modifying argument nominalization. On the other hand, in a construction like the man, who you love dearly, (is a real crook), it is likely that the denotation set of the modifying nominalization contains multiple entities, denoting all those that the addressee loves dearly. In this case, the denotation of the modifying nominalization identifies that of the head noun as an instance; namely, that the denotation of the head noun man is an instance of those that the addressee loves dearly. In both these cases, the modifying nominalizations identify the denotation of the head noun in terms of alternative and more elaborate ways of determining it, resulting in the commonly-held observation that a non-restrictive relative clause provides additional information about the denotation of the head noun.

The canonical RC constructions form a noun phrase with a head nominal and a modifying argument nominalization. Many languages, including English, however, allow the modifying nominalization to be separated from the NP containing the head noun, as in the examples below (see Fleck, this volume, for more examples).
(5-91) Tapiete (González 2005: 232)

\[\text{hama angu’ı nohe-ha} \quad [\text{wi-ro-po-ha-p i-wa}] \quad \text{hoka} \]

then drum pull out-imper 3-com-dance-imper-inst-nmlzr dem ha’e (cf. 5-9)

(s)he ‘then they take out the drum to dance, that is, (lit.) ‘…the drum with which they dance …’

(5-92) Harakmbut (Van Linden this volume)

\[\text{arakmbut-ta i’-uk-i} \quad [\text{henpu wa-mba-baeri}]\text{-ta} \]

person-acc 1sg-search-1.ind string.bag nmlzr-vrl-make-anim-acc ‘I am looking for the person who makes string bags.’

These are reminiscent of what Hale (1976) refers to as the adjoined relative clause in Australian languages, although it is not clear whether or not these and similar constructions in other parts of the world are “typically …separated from the main clause by a pause” (Hale 1976: 78).

Relevant to the interpretations of these “adjointed relative clauses” is the question of whether or not the modifying nominalizations themselves form a noun phrase and hence are referential. Our analysis of the canonical RC constructions, as shown in (5-90) above, juxtaposes the head noun and the modifying argument nominalization similar to an appositive construction of the type, John, the butcher. However, juxtaposition by itself does not mean that the two nominal constituents are in the appositive relation seen in the type given here that involves identification of the referent of one NP by another referential NP. Our RC analysis maintains that the modifying argument nominalization is not a noun phrase and hence, while it denotes an entity, it does not refer to a discourse entity, just as nouns in a noun compound such as \[[\text{goat}]_N \ [\text{cheese}]_N\] do not refer individually (or even collectively as a compound noun)—the entire noun phrase containing the compound, as in We ate \[[\text{goat}]_N \ [\text{cheese}]_N\]NP refers. Likewise, the nominalization structure identified as […]nmlz in this paper does not refer by itself; it only refers when it heads an NP.46

The discussion above is relevant in the interpretations of the two similar constructions in Spanish below.

46. Our field has seen many loose uses of the term “referential/referring expression” in reference to nouns and nominalizations (Croft 1991, Shibatani 2009, Cristofaro this volume) that ignore an important distinction between denotation and reference (see Section 4 on this distinction).
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(5-93) Spanish
   a. \([\text{El hombre} \ [\text{que está bailando}]]_{\text{NMLR}}\text{es mi tío.}\)  
      the man \text{the} \text{man} \text{NP} is dancing is my uncle  
      ‘The man who is dancing is my uncle.’
   b. \([\text{El hombre}]_{\text{NP}} \ [\text{el} \ [\text{que está bailando}]]_{\text{NMLR}}\text{es mi tío.}\)  
      the man \text{the} \text{NP} is dancing is my uncle  
      ‘The man, the one who is dancing, is my uncle.’

In (5-93b), the nominalization marked by the definite article forms a noun phrase and is therefore referential, while the counterpart in (5-93a) does not form an NP and does not refer.

Indeed, a contrast similar to that observed in the Spanish examples above is seen in a number of languages that have a special marker for an NP-use of grammatical nominalizations like the Spanish articles, e.g. \(\text{el}\) in (5-93b). Toba has demonstrative determiners that mark an NP-use of grammatical nominalizations (see Section 6). These demonstrative determiners, like the Spanish articles, do not mark nominalizations when they are in modification-use, as seen in (5-93a) above and (5-94a) below. Demonstrer marking in these languages, therefore, signals an NP-use of nominalizations, where they are playing a referential function. Observe the parallel pattern between (5-93) and (5-94).

(5-94) Toba (courtesy of Cristina Messineo)
   a. \(\text{ajem si-kjen} \ [\text{so} \ [\text{fijaGawa}] \ [\text{Ø-neta-ge da cako}]]_{\text{NMLZ}}\text{NP}\)  
      I 1A-greet \text{DD} \text{man} 3-be-DIR \text{DD Chaco}  
      ‘I greeted the man who lives in Chaco.’
   b. \(\text{ajem si-kjen} \ [\text{so} \ [\text{fijaGawa}]]_{\text{NP}} \ [\text{so} \ [\text{Ø-neta-ge da cako}]]_{\text{NMLZ}}\text{NP}\)  
      I 1A-greet \text{DD} \text{man} \text{DD} 3-be-DIR \text{DD Chaco}  
      ‘I greeted the man, the one who lives in Chaco.’

It is this kind of contrast that must be investigated in order to understand the true nature of the “adjoined” type seen in (5-91) and (5-92), or even the ordinary juxtaposed one, for that matter.

It is by now clear that the traditional term “relative clause” is a misnomer since what it designates is an argument nominalization in modification-use. The term can now be understood as a label for it. Such a label, however, is misleading since it suggests that nominalizations are clauses. We show below that there

47. As in the simple \(\text{John, the butcher}\) type, the constituency of the two juxtaposed NP’s is not clear; i.e. whether or not they form a larger NP constituent.

48. In Japanese, the referential appositive nominalization occurs after the head noun (e.g., \(\text{kono hon, [boku ga kinoo katta] no}\), ‘this book, the one that I bought yesterday’), while the ordinary RC is prenominal (\(\text{[boku ga kinoo kata] kono hon}\) ‘this book that I bought yesterday’).
is ample crosslinguistic evidence that nominalizations are not clauses, let alone sentences. But first a clearer understanding is in order of what sentences, clauses, and nominalizations are.

6. Sentences, clauses, and nominalizations

An early recognition of the use of grammatical nominalizations as noun modifiers is found in the context of Japanese grammar in Yoshio Yamada’s monumental grammar, *Nihonbunpōron* (Theory of Japanese Grammar) published in 1908. More recently, Matisoff (1972) recognized a connection between nominalizations and relative clauses in the Tibeto-Burman language Lahu. He also mentions similar morphological connections among nominalizations, relative clauses, and possessive constructions in Lahu, Mandarin Chinese, and Japanese (see Section 7). Matisoff’s observation has been followed by others specializing in Tibeto-Burman languages, such as DeLancey (1986) and Noonan (1997), and more recently by DeLancey (2002) and Noonan (2008).

Many other recent studies on nominalizations and relative clauses such as those contained in Yap, Grunow-Hårsta & Wrona (2011) and Comrie & Estrada-Fernández (2012) clearly recognize the use of nominalizations as relative clauses, but for some unclear reason and without any justification they continue to use the term “relative clause”, (i) as if some structures identifiable as relative clauses exist apart from argument nominalizations, but “in certain languages relativization is indistinct from nominalization” (Comrie & Thompson 1985/2007: 379) or (ii) as if nominalizations somehow turn into clauses under modification-use. Yamada (1908: 1462) simply states that we may call a grammatical nominalization used for noun modification an “adnominal clause” without offering the reason for it.

6.1 Tense and nominalization

Without clear definitions of clauses and sentences (and nominalizations, for that matter!) on the part of those who believe that relative clauses (our grammatical argument nominalizations) are clauses, it is difficult to know the true rationale underlying their belief about the clausehood of RCs. However, one observation that has been made is that RCs in some languages may stand as sentences; hence RCs are sentences that have been made dependent clauses by embedding them.49

49. Rice (1989: 25), in her otherwise excellent grammar, tells us that “[a] relative clause is a sentence that modifies a noun”. Compare this with Nevis, Pesetsky & Rodrigues’s (2009: 366) characterization of event nominalizations: “a verb may merge with a sentence, as in *Mary thinks*
Such a possibility arises when the predicate in an RC has a tensed verb or more broadly a finite verbal form associated with sentences. For example, Comrie and Horie (1995: 68) tell us that “[w]hat precedes the head noun [as in (6-1a) below, for example] is a well-formed sentence in its own right”, as can be seen from the fact that it can stand as a sentence; see (6-1b).

(6-1) Japanese
a. \([\emptyset_1 \text{kino} \emptyset_2 \text{katta}] \text{hon}\]
yesterday bought book
‘the book that (I) bought yesterday’
b. \(\emptyset_1 \text{kinoo} \emptyset_2 \text{katta}.\)
yesterday bought
‘(I) bought (it) yesterday.’
(As an answer to the question “When did you buy the book?”)

Comrie and Horie are equating the gaps found in the RC in (6-1a) with the anaphoric gaps found in sentence (6-1b). This, however, is a mistake. Anaphoric gaps can be filled by full noun phrases, albeit perhaps redundantly, but one of the gaps in the RC/argument nominalization in (6-1a) cannot. Compare (6-1) with (6-2) below:

(6-2) a. \([\emptyset_1/\text{boku ga} \emptyset_2/\text{*sono hon o katta}] \text{hon}\]
I nom yesterday that book acc bought book
‘the book [that I bought \(\emptyset/\text{that book}\)]’
b. \(\emptyset_1/\text{Boku wa} \emptyset_2/\text{sono hon o katta}.\)
I top yesterday that book acc bought
‘I bought that book yesterday.’

(6-2a) is as bad as its English translation with the full noun phrase in object position. In other words, the two gaps in RC (6-1a) are different from the two gaps in sentence (6-1b). Object argument nominalizations must have an obligatory gap (\(\emptyset_2\)) in object position in both English and Japanese, while the latter may contain an anaphoric gap in other positions.\(^50\) Clauses and sentences, on the other hand, have no such constraint. Argument nominalizations are thus different from clauses and sentences in both English and Japanese.\(^51\)

\[\text{that the world is round}\]” and “…a noun can merge with a sentence, as it does in \((\text{the}) \text{ claim [that the world is round]}\),…” (Emphasis added).

50. Japanese, as in some other languages, allow a resumptive pronoun in a position lower in the grammatical relation hierarchy.

51. Comrie and Horie (1995) recognizes this difference in footnote 5, page 75, but does not deal with this most crucial issue in comparing the structures of RCs/argument nominalizations and
A similar conclusion obtains with the Mayan language K'ichee' spoken in Guatemala, whose argument nominalizations modifying a noun qua RCs contain a finite verb form and appear to be able to stand as sentences, as shown in (6-3b) below.

(6-3) K’ichee’ (Larson & Norman 1979: 357; the grammaticality judgement courtesy of Telma Can Pixabaj)

a. *lee ixoq lee [x-Ø-u-ch’ay leed achih] the woman NMLZR ASP-3SG.ABS-3SG.ERG-hit the man ‘the woman whom the man hit/the woman who hit the man’

b. x-Ø-u-ch’ay lee achih ASP-3SG.ABS-3SG.ERG-hit the man ‘S/he hit the man/The man hit him/her.’

However, just like the case of Japanese above, nominalizations qua RCs are different from sentences. The latter can have a full array of arguments appearing as full noun phrases, while the former must contain a gap. Observe:

(6-4) a. lee ixoq lee [x-Ø-u-ch’ay lee achih] the woman NMLZR ASP-3SG.ABS-3SG.ERG-hit

*lee ixoq/Ø lee achih
the woman/Ø the man
‘the woman whom [the man hit *the woman/Ø]’ or ‘the woman who [‘the woman/Ø hit the man’]’

b. x-Ø-u-ch’ay lee ixoq lee achih
ASP-3SG.ABS-3SG.ERG-hit the woman the man
‘The woman hit the man/the man hit the woman.’

The difference between argument nominalizations qua RCs and clauses/sentences seen here also obtains in those languages that may contain a pronoun instead of a gap in argument nominalizations/RCs. Thus, the pronoun in subject position of a subject nominalization cannot be replaced by a full noun, as shown in (6-5b) below, which is just as bad as its English translation.

(6-5) Thai

a. thɔɔ mây khuan kin yaa [thii man mòtaayûʔ] 2SG not should eat medicine NMLZR 3SG expire (lit.) ‘You should not take the medicine which it has expired.’

b. *thɔɔ mây khuan kin yaa [thii yaa mòtaayûʔ] 2SG not should eat medicine NMLZR medicine expire (lit.) ‘*You should not take the medicine [which the medicine has expired].’

(complement) clauses/sentences. Other related papers by Comrie (Comrie 1996 and 1998a, b) repeat similar views about Japanese and other languages without even mentioning this issue.
Similar examples can be adduced from a diverse array of languages whose argument nominalizations contain verbal forms similar to those occurring in sentences, with tense and other finite features or without any of them, as in isolating Asian languages like Thai above.

The reluctance to recognize these nominalizations as such is rooted in the fact that they may contain formal finite features such as tense, aspect, and person marking, characteristics of sentences as in the examples above. However, there is nothing that prevents nominalizations from having these features since the information they carry can be highly valuable in distinguishing types of entities they denote. For example, what-nominalizations in English make a crucial difference in what they denote depending on the tense information they contain; e.g. *what I was buying* vs. *what I am buying*; also cf. *John's purchasing of a house last year* vs. *John's purchasing of a house next year*. Indeed, in many languages nominalizing morphology itself may incorporate tense information (and/or aspectual, as well as evidential values in some languages), as shown in (6-6)–(6-8), or may allow a separate tense expression within nominalized structures, as in the Oceanic examples given in (6-9) and (6-10) below.52

(6-6) Korean

a. \[cikum \text{ pap-ul \ mek-nun}] \text{ kes}
   \text{now meal-ACC eat-PRS.NMLZR NPM}
   'one who is eating a meal now'
   (cf. \[cikum \text{ pap-ul \ mek-nun}] \text{ ai} 'a child who is eating a meal now')

b. \[ecey \text{ pap-ul \ mek-un}] \text{ kes}
   \text{yesterday meal-ACC eat-PST.NMLZR NPM}
   'one who ate a meal yesterday'

c. \[pap-ul \text{ mek-ul}] \text{ kes}
   \text{meal-ACC eat-FUT.NMLZR NPM}
   'one who will eat a meal'

(6-7) Hixkaryana (Derbyshire 1999: 48–49)

a. Event/Action nominalizer-Past tense: -thiri
   \text{i-wanota-thiri \ komo}
   3-sing-AC.NMLZR COL
   'their singing (in the past)'

52. See Fleck (this volume), Peña (this volume), Valle & Zariquiey (this volume), and Jones & Jones (1991) for additional cases of time-bound nominalization morphology in Matses, Wampis, and Barasano. Also compare these cases with so-called present and past participles in English forms, a *breaking chair*/*a broken chair*. 
b. Nominalizer of the S (Protagonist of intransitive event)/O (Patientive protagonist of transitive event)-Past tense: -saho
   i-manho-saho  uro/omoro/moki
   IMPERS-dance-S.NMLZR 1/2/3PRO
   ‘I (am)/you (are)/he (is) the one who danced.’

(6-8) Imbabura Quechua (Cole 1982)
   a. [Marya Ø riku-shka\textsuperscript{53}] runa
      Maria see-PST.NMLZR man
      ‘the man whom Maria saw’
   b. [Marya Ø riru-k] runa
      Maria see-PRS.NMLZR man
      ‘the man whom Maria sees’
   c. [Marya Ø riku-na] runa
      Maria see-FUT.NMLZR man
      ‘the man whom Maria will see’

(6-9) Xârâcùù (Oceanic; Moyse-Faurie 2016: 182)
   È kê pwi [êê-mwata na rê anyâå]
   3.sg eat.tubes banana NMLZR-grate PST POSS mommy
   ‘He is eating bananas which have been chopped up by his mother.’

(6-10) Marquesan (Oceanic; Moyse-Faurie 2016)
   [Te i ite-tina na tunane tata eka te tihe te kui
    SPEC PST see-NMLZR PAUC brother nearly reach SPEC COME SPEC mother
    i una], atahi kokoti na tunane te ouoho no Hina.
    LOC TOP then cut PAUC brothers SPEC hair POSS Hina
   ‘When the brother saw that the mother had nearly reached the top, (then)
   they cut Hina’s hair.’

Tense is intimately connected with a sentence because the latter asserts the truth
of a predication made by a clause as obtaining at a specific time. Since nominal-
izations do not perform this kind of illocutionary function, they typically lack an
expression of tense. On the other hand, tense indication one way or another adds
some vital information about what is denoted by nominalizations. There is thus no
need to assume that nominalizations cannot be marked for tense, and we should
not uncritically assume that tense-marked structures are clauses or sentences.

\textsuperscript{53} The Imbabura -shka corresponds to the object argument nominalizer/event nominal-
izer -sqa, and -k to the subject argument nominalizer -q in Bolivian Quechua. The connec-
tion between object argument nominalizer and past tense is seen elsewhere, e.g. Kakataibo
(Zariquey 2011).
6.2 Functional definitions of clauses, sentences, and nominalizations

Those who believe that argument nominalizations (used in RC constructions) and other types of grammatical nominalizations are clauses/sentences are victims of the formal orientation in linguistics that attempts to characterize the nature of grammatical constructions in terms of formal properties such as internal structural properties. As seen above, grammatical nominalizations (partially) share internal structures with clauses and sentences. But these structure-internal formal properties are like the skeletal structures that a roast turkey shares with a live one. Just as a roast turkey and a live bird are functionally very different and are accordingly treated differently (e.g. we do not keep them in the same cage), grammatical constructions such as clauses, sentences, and nominalizations must similarly be defined functionally and distinguished according to the functions they perform. Shibatani (2017, 2018a) offers the functional definitions of clauses, sentences, and nominalizations along the following lines, in terms of different kinds of speech act they perform:

- **Clauses predicate**: By uttering a structure like [John is honest] as a clause, a speaker ascribes the relational property denoted by a predicate phrase to the referent of the subject noun phrase.

- **Sentences perform illocutionary acts**: By uttering a structure like [John is honest] as a declarative sentence, a speaker asserts the truth of the predication made by the clause of the same structure. By uttering a structure like [Is John honest] as an interrogative sentence, a speaker asks whether the predication made by the clause [John is honest] is true or not.

- **Nominalizations denote** (things and thing-like entity concepts): By uttering a structure like [John is honest] as a nominalization, a speaker evokes and establishes a form-meaning connection between the structure and a metonymically motivated meaning, e.g. a fact, related to a state-of-affairs (or broadly an event).

In other words, a clause is a grammatical structure associated with a speech act of ascribing a verbal property to the referent of a subject nominal. Sentences, on the other hand, perform different kinds of speech act, namely illocutionary acts, such as asserting that the predication made by a clause is true (declarative sentences), questioning whether or not the predication is true (yes-no questions), ordering (imperative sentences), warning, promising, etc. Notice that predication and assertion are two distinct types of speech act, as clearly shown in the case of yes-no questions. Nominalization structures neither predicate nor assert; they instead denote. That is, by uttering a noun or nominalization, a speaker evokes a mental connection between a string of sounds with an entity concept. Many such mental connections are permanent, as in the case of nouns and lexical nominalizations,
while grammatical nominalizations evoke nonce connections between the forms and concepts.

Besides the structural properties that nominalizations may share with clauses and sentences, as in our examples here, meaning similarities among them is likely another reason that many believe that nominalizations are clauses or sentences. For one thing, nominalizations share meaning similarities with clauses and sentences in terms of presuppositions that they are associated with. The event nominalization [(that) John recklessly shoots trespassers] in a sentence like Bill knows (that) John recklessly shoots trespassers presupposes the proposition “John recklessly shoots trespassers”. Likewise, the argument nominalization [who John shot] qua a relative clause as in I knew the man who John shot presupposes the proposition “John shot someone/something”. How one arrives at these presuppositions from the nominalization structures is an interesting and difficult question. One should not presume that meaning resides in structures in a straightforward manner. A presuppositional proposition can also be derived from a sentence that does not contain a clausal structure corresponding to the form of a presupposition in question; e.g. John is a good cook presupposes “John cooks”; The nurse attended the crying baby presupposes “The baby was crying”.

Perhaps a more fundamental issue here is a distinction between linguistic meaning and propositional meaning, the latter of which obtains when a structure (a linguistic form, a series of speech sounds) is used in performing speech acts of predicatation and assertion. A structure can have linguistic meaning apart from these speech acts, as in the case of those metalinguistic statements used as examples in linguistics discussions, such as John is honest, John sent a Christmas gift to his grandmother, which have full and coherent meanings similar to actual clauses and sentences, while they lack a referential property and truth value, i.e. propositional meanings of clauses and sentences. In short, a nominalization may share meaning similarities with a clause/sentence at the level of linguistic meaning, however such a thing is arrived at. And it is at this level that the internal structure of nominalizations plays an important role in determining their meanings and their similarity to those of clauses/sentences.

The event nominalization of the form [that [John drank beer excessively]] accordingly shows a high resemblance in linguistic meaning to the clause/sentence John drank beer excessively, whereas those that show only partial structural similarities, such as [John’s drinking beer excessively], [drinking beer excessively],

54. We have not made much progress on this from the early phase of Generative Grammar, where an attempt was made to arrive at linguistic meaning by an algorithm that compositionally built up the meaning on the basis of phrase structures/markers. Fillmore’s Case Grammar was also an attempt to capture meaning similarities across different syntax structures.
([drinking excessively], and [drinking], share partial meaning similarities to the relevant clause/sentence according to their structural complexity. The difference in structural complexity of this type correlates with the generality in meaning, from the most specific for [that [John drank beer excessively]] to the least specific for [drinking], but it does not make one nominalization functionally “more/less nominalized” than others.

Being nominal, nominalizations may head an NP and function as arguments of clauses and sentences. They do not stand alone like sentences in their capacity as nominal structures. However, nominalizations may become clauses and sentences when they become used to perform illocutionary acts, just as a noun can be used as a sentence issuing a warning (Fire!), for example (see next section). Conversely, sentences/clauses do not function as NP arguments. The only case in which they function as arguments is when used as a direct quotation; e.g. Muhammad Ali said/boasted/wrote, “I am the greatest of all!” In this way, grammatical constructions—sublexical morphemes, words, as well as larger phrasal units—are defined in terms of their functions, not by their formal or meaning similarities to other structures, though these provide supporting evidence for treating alike structures bearing the same function.

From our perspective, the terms “clause” and “nominalization” are mutually exclusive. The terms “nominalized clause” and “clausal nominalization” used widely in the field make sense only in one reading, namely in referring to nominalizations that share structural similarities with clauses but which are not clauses, not in the sense of the nominalizations that are clauses, which we believe do not exist. In the face of the proposed terms such as “grammatical nominalization” and “event nominalization” in this paper, it is not clear whether ambiguous terms such as “clausal nominalization” and “nominalized clause” have any theoretical status, besides the difficulty in precisely determining the degree of structural similarity that a nominalization must have in order for it to qualify as a clause.

6.3 Insubordination

A discussion of the distinction between clauses and sentences, on the one hand, and nominalizations, on the other, cannot be complete without touching on the use and development of nominalizations as sentences. The term insubordination or desubordination refers to a phenomenon in which a dependent structure, or a structure that does not stand as a sentence by itself, comes to be used as a stand-alone sentence (see Evans 2007). There are several clear cases in which nominalizations, which usually do not stand as sentences, develop into sentences. We discuss here two such cases; one, where event nominalizations as a whole get reanalyzed as sentences, eventually replacing sentences marked by a finite verb, and the other, where a sentence develops out of nominalizations combined with “auxiliary verbs” via a dropping of the latter.
In modern central dialects of Japanese, including Tokyo Japanese, the finite verb (known as *shūshi-kei* ‘conclusive form’) and the nominalized verb form (known as *rentai-kei* ‘adnominal form’), which were distinct in Old Japanese (8th C) for many verbs, are identical, except for the copula, whose finite form is *da* and the nominalized form *na*. The merger of the finite and the nominalized verb form resulted from the use of nominalizations as sentences, which had been seen from the time of Old Japanese through Early Middle Japanese, and which was completed during the Late Middle Japanese period (13th–17th C). The following, from the *Genji Monogatari* (*The Tale of Genji*) first published in 1008, is a well-known example of the use of an event nominalization as a sentence.

(6-11) *Suzume no ko o Inuki ga nigasi-turu.*

sparrow GEN child ACC Inuki GEN let.go-PERF.NMLZR
(lit.) ‘Inuki’s letting my baby sparrow go–(shucks!).’

The perfective ending in the above example is in nominalized form, whose finite form is *-tu*, which would end usual declarative sentences. The use of the event nominalization as a sentence above was made possible because it performed the illocutionary act of evincing a feeling of disgust, regret, lamentation, or surprise. When the use of forms like this continued to expand, their illocutionary forces eroded, and they have eventually been reanalyzed as ordinary declarative sentences, replacing those that ended in the finite verb form. This change brought about a reinterpretation of the genitive form modifying a nominalization (*Inuki ga*) as a subject. Stand-alone event nominalizations of the above type with a special illocutionary force have been reported elsewhere, where nominalizers involved are analyzed as a “stance-marker” (see Yap & Grunow-Hårsta (2010)).

Japanese nominalizations, being nominal, cannot predicate over a subject referent unless they combine with a copula or some tense-carrying “auxiliary verbs”, many of which are grammaticalized versions of verbs, such as *suru* ‘do’, *yaru* ‘give’, *iru* ‘exist/be’ and their honorific variants. Compare the following forms.

(6-12) Tokyo Japanese

a. *Taroo ga tegami o kai-ta.*

Taro NOM letter ACC write-PST
‘Taro wrote a letter.’

b. *Taroo ga tegami o kak-i nasat-ta.*

Taro NOM letter ACC write-NMLZR do.HON-PST
‘Taro wrote (HON) a letter.’

---

55. There is a good possibility that nominalization chains discussed earlier are precursors of these constructions. See (5-56) and (5-57).
Sentence (6-12b) contains the nominalized form *kak-i* ‘writing’, which by itself cannot predicate. It can, however, combine with the honorific auxiliary *nasar- ‘do’, which inflects for tense. *Nasaru* has the somewhat archaic-sounding imperative form *nasare*, which in Western Japan dialects is pronounced as *nahare*, as in (6-13a), below. Now, Western Japan dialects, but not Tokyo Japanese, may drop this imperative auxiliary and let the form ending in nominalization carry an imperative force; hence turning a nominalization into a sentence, as in (6-13b).

(6-13) Western Japan dialects (Osaka, Kyoto)

a. *Tegami* o *kak-i* *nahare*.
   letter ACC write-NMLZR do.HON
   ‘Write a letter (honorific)/(lit.) Do writing a letter (honorific).’

b. *Tegami* o *kak-i!*
   letter ACC write-NMLZR
   ‘Write a letter (plain)!’

Tokyo Japanese, on the other hand, allows a dropping of a related polite imperative auxiliary in the form *nasai* ‘do (HON)’, when the nominalized form is honorified by the use of the prefix *o-*, as seen below.56

(6-14) Tokyo Japanese

a. *Tegami* o *o-kak-i* *nasai*.
   letter ACC HON-write-NMLZR do.POL
   ‘Write a letter (polite).’

b. *Tegami* o *o-kak-i!*
   letter ACC HON-write-NMLZR
   ‘Write a letter (polite)!’

A more widespread pattern in Japanese dialects involves combinations of the conjunction-marked nominalization *V-i=te*, which also combines with a variety of inflecting auxiliaries, including the grammaticalized forms of the verbs of giving marking a benefactive sentence, as in (6-15a) below. The dropping of the auxiliary verbs results in nominalization-based imperatives, as in (6-15b).

(6-15) Tokyo Japanese

a. *Tegami* o *ka-i=te* (< *kak-i=te*) *kure/kudasai*.
   letter ACC write-NMLZR=CON give.me/give.me.POL
   ‘Write a letter for me/Write me a letter.’

b. *Tegami* o *ka-i=te!*
   letter ACC write-NMLZR=CON
   ‘Write a letter for me/Write me a letter!’

56. Cf. *tegami* ‘letter’ : *o-tegami* ‘letter (honorific)’. 
Finally, Japanese also allows a dropping of the copula *da* and its honorific variant *desu*, as well as the interrogative particle *ka*, allowing an event nominalization marked by the NP-use marker *no* to stand as a sentence. Compare;

(6-16) Tokyo Japanese
   a. *Taro* ga *kinoo* kita no *desu* ka?
      Taro  NOM yesterday came NPM COP Q
      (lit.) ‘It is that Taro came yesterday?’ ‘Did Taro come yesterday?’
   b. *Taro* ga *kinoo* kita no?
      Taro  NOM yesterday came NPM
      (lit.) ‘It is that Taro came yesterday?’ ‘Did Taro come yesterday?’

Since the copula *da/desu* can drop by itself, a declarative sentence identical in form to the interrogative sentence (6-17b) also obtains; the interrogative form (6-16b) is differentiated from the declarative one (6-17b) by a rising intonation.

(6-17) Tokyo Japanese
   a. *Taro* ga *kinoo* kita no *da/desu*.
      Taro  NOM yesterday came NPM COP/COP.POL
      (lit.) ‘It is that Taro came yesterday.’ ‘Taro came yesterday.’
   b. *Taro* ga *kinoo* kita no.
      Taro  NOM yesterday came NPM
      (lit.) ‘It is that Taro came yesterday.’ ‘Taro came yesterday.’

Like the other nominalization-based sentences, (6-17b) is usable only as a response to a question such as (6-16a), (6-16b), or in response to someone demanding an explanation for one’s action. The sentential use of nominalizations, in other words, is most prevalent in face-to-face inter-personal situations, where the context allows the hearer to figure out the illocutionary force intended by the speaker.

The discussion on the insubordination phenomenon in Japanese above bears on the issues surrounding the treatment of the Pirahã nominalization raised by Everett (2005, 2009). In these papers Everett reverses his earlier nominalization treatment (Everett 1986) of the *sai*-marked structures seen below.

(6-18) Pirahã (Mura-Matanawi; Brazil)
   a. *hi ob-áaxáí* [kahai kai-sai]
      see/know-INTNS arrow make-NMLZR
      ‘He really knows how to make arrows.’
   b. *kóxoí soxóá xibíih-i-háí* [tiobáahai biio kai-sai]
      Kóxoí already order-PROX-REL.CERT child grass do-NMLZR
      ‘Kóxoí already ordered the child to cut the grass.’
Everett (2005, 2009) now identifies the suffix *-sai* as a marker of old information, and reanalyzes a sentence like (6-18a) as a paratactic coordination, as in the manner below.

(6-19) (Hi) xob-áaxáí. (Hi) kahaí kai-sai.
(3) see-well (3) arrow make-old.info
‘He is really smart/very talented. (That is with respect to the fact that) he makes arrows.’ (Everett 2009: 410)

Everett’s motivation for his reanalysis of the Pirahã nominalizer *-sai* as an old information marker lies in his observations and conclusion below:

“The second verb above [(6-19)], ‘to make’, is a bare root followed by *-sai*. This looks like nominalization until we see that: (i) the verb can take a full range of inflection …; (ii) that the sentence in which *-sai* appears can also appear as a main clause [(6-20) below].”

“If both clauses refer to topical information, both can bear the *-sai* suffix. If *-sai* were a nominalizer, however, we would not expect it to appear on both clauses since, presumably, a nominalized clause would not be a stand-alone sentence (cf. *John running the store*, *Rome’s destruction of Carthage*).” (2009: 410)

(6-20) a. Kóhoi xob-áaxáí xáagí-sai.
name see-well permanent:to.be-old.info
‘Kóhoi really knows his stuff.’
b. Kóhoi hi kahaí kai-b-íigí-sai.
name 3 arrow make-move:down-cont-old.info
‘He is finishing making arrows.’ (Everett’s glosses)

Clearly Everett does not know that grammatical nominalizations may have fully inflected verb forms (e.g. [that [John might have been arrested]]). Neither is he aware of the phenomenon of insubordination, which by itself does not invalidate a nominalization analysis.\(^57\)

6.4 Evidence that nominalizations are not clauses or sentences

There are some compelling pieces of evidence pointing to the nominal nature of grammatical nominalizations that help distinguish them from clauses and sentences. Below we examine the two quintessentially nominal phenomena of plural and classifier marking.

\(^{57}\) The reason that *sai*-marked nominalizations are associated with the notion of old information has likely to do with their associated presuppositions. See the contributions to this volume by Bruil, Gipper & Yap, Machado & Peña for additional cases and discussions of insubordination.
6.4.1 Plural marking

Languages that have plural marking on nouns may mark grammatical argument nominalizations similarly since both may denote countable entities. Observe the following Bolivian Quechua forms.

(6-21) Bolivian Quechua

a. wasi ‘house’: wasi-kuna ‘houses’
   llank’a-q ‘worker’: llank’a-q-kuna ‘workers’

b. [wallpa-ta wayk’u-q]-kuna
   chicken-ACC cook-SUB.NMLZR-PL
   ‘ones who are cooking a chicken’

c. [[wallpa-ta wayk’u-q]NM_LZ-kuna] warmi-kuna
   chicken-ACC cook-SUB.NMLZR-PL woman-PL
   ‘women who are cooking a chicken’

The examples (6-21b.c) show that the subject grammatical nominalization involved plays a denoting function, just like a simple noun wasi ‘house’ in (6-21a), rather than the predication or the assertion function of a clause and a sentence. Notice, however, that a Quechua sentence, as in some other languages, may contain a verb marking plurality of an NP referent within a sentence, as in the following sentence.

(6-22) Waki-n runa humu-n-ku.
   some-3 man come-3-PL
   ‘Some of the men come.’

Crucially, the plural morpheme marking verbs differs from that marking nominals, although there is an obvious similarity in form.

Similar plural marking of grammatical nominalizations is seen in a fair number of languages, as the following data show.

(6-23) Capanawa (Panoan; Peru; Loos 1999)

[ʔəá tsaʔot-ai]NM_LZ -bo his-i
   there sitPRS -PL see-IMER
   ‘Look at those (who are) sitting over there.’

Cf. [ʔənɨ hiwi mebi taʔpat-ai]NM_LZ teʔpan anin tiso honi
   big tree branch bifurcatePRS fork LOC monkey hidePRS
   ‘A monkey is hiding in the fork of a branch that bifurcates from a large tree.’

(6-24) Nheengatu (Tupí-Guaraní; Brazil; Cruz 2011 and p.c.)

a. re-su re-mu-tawari kau [re-yu-mu-kuaku]
   2SG.A-go 2SG.A-CAUS-tobacco DEM 2SG.A-R/R-CAUS-be.fasting
   wa]=ita u-mbau ara
   NMLZR=PL 3SG.A-eat PROS
   ‘You are going to bless those whom you made fast.’
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(6-25) Yaqui (Alvarez 2012 and p.c.)

a. \[\text{[in } \text{j} \text{inu-ka-root-}m\text{ sikili} \]
   \hspace{1cm} 1SG.GEN buy-PERF-NMLZR-PL red
   ‘Ones I bought are red.’

b. \[\text{U-me bisikleeta-m [in } \text{j} \text{inu-ka-}m\text{ sikili} \]
   \hspace{1cm} DET-PL bicycle-PL 1SG.GEN buy-PERF-NMLZR-PL red
   ‘The bicycles that I bought are red.’

(6-26) Salve (Rice 1989: 83)

\[\text{[niwá kedaw’i] i } \text{ke gogháyeyida} \]
\hspace{1cm} long 3PL.sat NMLZR PL 1SG.SAW.3PL
‘I met ones who stayed a long time.’ (Hare dialect)

Turkish also allows plural marking on grammatical argument nominalizations, but does not permit doubling of plural marking on both the head noun and the modifying nominalization, as in (6-27c) below.

(6-27) Turkish (Göksel & Kerslake 2005: 449 and Yu Kuribayashi p.c.)

a. \[\text{Opera-yı sev-me-yen}_{\text{NMLZ-}ler-e } \text{şasyorum}. \]
   \hspace{1cm} opera-ACC like-NEG-NMLZR-PL-DAT surprised.1SG
   ‘I am surprised at those who don’t like opera.’

b. \[\text{[[Opera-yı sev-me-yen}_{\text{NMLZ, kişi-ler}-e \text{ şasyorum}. \]
   \hspace{1cm} opera-ACC like-NEG-NMLZR PERSON-PL-DAT surprised.1SG
   ‘I am surprised at the people who don’t like opera.’

c. *\[\text{[[Opera-yı sev-me-yen}_{\text{NMLZ-ler}-e \text{ şasyorum}. \]
   \hspace{1cm} opera-ACC like-NEG-NMLZR PERSON-PL-DAT surprised.1SG
   ‘I am surprised at the people who don’t like opera.’

Tapiete grammatical argument nominalizations, in addition to plural marking, show another nominal feature, foreign to clauses and sentences, namely, diminutive marking, as below.

(6-28) Tapiete (Golzárez 2005 and Coccine 2008)

a. \[\text{o-che-wa-reta} \]
   \hspace{1cm} 3AC-sleep-NMLZR-PL
   ‘(the ones) who are sleeping’

b. \[\text{karai-re } \text{[tumpa i-ñe’e} \text{ mbe’u i-a-reta]}. \]
   \hspace{1cm} white.man-PL god 3.POSS-language tell be-NMLZR-PL
   ‘the gringoes (white men) who were announcing (reciting) the Bible’
6.4.2 Classifier marking

The Piapoco data bring us to the next nominal feature that reflects the entity-denoting property of nominalizations, namely classifiers. Japanese numeral classifiers, mostly Chinese loans, occur in several syntactic positions. Two common patterns are shown below, where the numeral classifier san-satu [three-clf.BOUND] ‘three bound (things)’ occurs prenominally (6-31a) and as an adverb away from the modified noun (6-31b).
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(6-31) a. Ken wa san-satu no hon o kinoo motte kita.
   Ken TOP three-CLF GEN book ACC yesterday carry.GER came
   ‘Ken brought three books yesterday.’

b. Ken wa hon o kinoo san-satu motte kita.
   Ken TOP book ACC yesterday three-CLF carry.GER came
   ‘Ken brought three books yesterday.’

Grammatical nominalizations in Japanese do not seem to readily allow prenominal numeral classifiers, but they can be quantified by adverbial numeral classifiers, indicating that grammatical nominalizations denote entities rather than predicate or assert like clauses and sentences.

(6-32) a. *Boku wa [san-satu no [Ken ga motte kita]_{nmzlz} no o
   I TOP three-CLF GEN Ken NOM carry.GER came NM ACC
   kinoo yonda.
   yesterday read
   (lit.) ‘I read yesterday three of what John brought.’

b. Boku wa [Ken ga motte kita]_{nmzlz} no o kinoo
   I TOP Ken NOM carry.CON came npm ACC yesterday
   san-satu yonda.
   three-CLF read
   ‘I read yesterday three of what Ken brought.’

It is interesting to notice that the adverbial quantifier in (6-32b) has the partitive interpretation of reading three of what Ken brought, rather than quantifying what is denoted by the grammatical nominalization. But the point is that the choice of an adverbial quantifier is determined by the denotation of the grammatical nominalization, as the comparison between (6-32b) and the following clearly shows.

(6-33) Boku wa [Ken ga motte kita]_{nmzlz} no o kinoo
   I TOP Ken NOM carry.CON came npm ACC yesterday
   san-bon nonda.
   three-CLF drank
   ‘I drank yesterday three (bottles) of what Ken brought.’

The choice of adverbial classifiers indicates different types of things that Ken brought. The use of satu in (6-32b) indicates that what Ken brought were books or book-like bound materials, while the use of hon/bon in (6–33) indicates that what Ken brought were contained in cylindrical containers such as bottles.

Interestingly Chinese allows the pattern in (6-34a) disfavored in Japanese. Observe:
(6-34) Mandarin Chinese
a. sān-běn shū  
   three-clf book  
   ‘three books’

b. sān-zhī niǎo  
   three-clf bird  
   ‘three birds’

   three-clf I buy-NMLZR very expensive  
   (lit.) ‘Three what [books] I bought were very expensive.’  
   Cf. [wō māi-de]₁₉₇₁ shū  
   I buy-NMLZR book  
   ‘book that I bought’

d. sān-zhī [wō māi-de]₁₉₇₁  
   three-clf I buy-NMLZR  
   (lit.) ‘three what [animals] I bought’

Forms (6-34c, d) show that the grammatical nominalizations wō māi-de ‘what I bought’ may denote a variety of things evoked by this structure, and depending on what they actually denote, different classifiers are chosen in quantifying the denoted objects, such as books or book-like materials as in (6-34c) and animals as in (6-34d).

That grammatical nominalizations denote, rather than predicate or assert, is also clearly seen from the use of classifiers in Thai, which allows optional classifier marking of grammatical argument nominalizations. Observe.

(6-35) Thai (courtesy of Kingkarn Thepkanjana)

a. khruu [lăaj khon]  
   teacher many CLF.PERSON  
   ‘many teachers’

b. māa [sìi tua]  
   dog four CLF.BODY  
   ‘four dogs’

c. (chán chòp) [[thìi khwêe nay tều]₁₉₇₁]NP  
   I like NMLZR hang in closet  
   (I like) the one hanging in the closet.  
   (Answer to the question “Which skirt do you like?”)

c’. (chán chòp) [tua [thìi khwêe nay t Morrow]₁₉₇₁]NP  
   I like CLF NMLZR hang in closet  
   (I like) the one hanging in the closet.  

d. (chán chòp) [kràprooŋ [tua [thìi khwêe nay t Morrow]₁₉₇₁]]NP  
   I like skirt CLF NMLZR hang in closet  
   (I like) the skirt hanging in the closet.”
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

e. *Tua [krâproŋ khwēe nay tâu]
   CLF skirt    hang in closet
   'A skirt hangs in the closet.'

Notice that a clause/sentence is never marked by a classifier, as indicated by the ungrammatical form in (6-35e) above.

While classifier marking is largely optional in Thai in the sense that relevant structures may or may not be marked by a classifier depending on how specific one wants to be with regard to what a nominalization denotes, some languages use classifiers, instead of a general nominalizer, in marking nominalizations. Cantonese is one such language, in which grammatical nominalizations may be marked by classifiers, the choice of which depends on what they denote.\(^{58}\)

a. [nī ɗi] yū
   this CLF fish
   'this fish'
b. [sām go] hohksāang
   three CLF student
   'three students’
   we in France eat CLF food quite good-eat NM
   'The food that we ate in France was pretty good.'
d. [[[Gāau léih tāahn kāhm] go] go]?
   teach you play piano that CLF
   'The one who teaches you (to play the) piano?'

Asia is not the only area in which nominalizing classifiers occur. The Amazon Basin is another area where classifiers play important grammatical roles, including their use as numeral classifiers and for marking grammatical nominalizations, again underscoring the point that grammatical nominalizations denote entities, which can be classified according to their nature. Observe the following data from Bora, where the classifier hà marks an argument nominalization denoting an object like a shelter or with a sheltering function and kpà a slab-like object.

(6-37) Bora (Witotoan; Colombia, Peru, Brazil; Thiesen & Weber 2012)
a. ó-a-i-thù̂m-i-[a'i-β̆ei]-hà
   I see-<t>    burn-sIn-<shelter>
   'I saw a house that was burning.' (lit.) 'I saw one (shelter-like thing) that was burning.'

58. The nominalizing function of classifiers is discussed fully in Section 8.
b. \( \ddot{o}k^{h}\ddot{e} \ t^{h} \dddot{a}:k^{h} \ddot{u}\) [\(\dddot{e}:\)-kp\(\ddot{a}\)] [\(\dddot{e}:h\ddot{a}\)] t\(\dddot{f}\ddot{i}:\dddot{m}\ddot{e} \ddot{f}\ddot{c}\) \\
1-objAn you.imp.give that-<slab> that-<shelter> below \\
k\(h\ddot{a}\)-kp\(\ddot{a}\)\(\dddot{b}\ddot{u}\) be-<slab>-thm

'(Give me the plank that is under the house.' (lit.) 'Give me the (slab-like thing) one (slab-like thing) which is under the (shelter-like thing).')

Like grammatical nominalizations in other languages, these nominalizations marked by classifiers also have a modification-use. First observe (6-38a) below, in which a gender-based classifier marks argument nominalizations, a phenomenon very common among Amazonian languages. Argument nominalizations like this can modify a head noun, as in (6-38b).

(6-38) a. [\(\dddot{O} \ h\dddot{o}\dddot{a}:k^{h}\ddot{e} \ u\dddot{r}k\ddot{p}\ddot{a}:p\dot{o}\)]::p\(\ddot{e}\) \(t^{h}:\dddot{a}:\dddot{\ddot{r}}\ddot{i}\) (379–380) \\
John-objAn teach-<sg.msc> came-<ct>

'(The one-msc) who taught John came.'

b. \(\dddot{o} \dddot{a}:\dddot{t}^{h}:\dddot{u}\dddot{m}\dddot{i}::p\ddot{e}:k\ddot{\ddot{h}}\ddot{e}\) [\(\dddot{O} \ \dddot{o}k^{h}\ddot{e} \ i^{\ddot{o}:t\ddot{\ddot{\ddot{o}}}}::p\ddot{e}\)-kh\(\ddot{e}\) (381) \\
I see-<ct> dog-<sg.msc> I-objAn bite-<sg.msc>-objAn

'I see the dog that bit me.

The Bora patterns above may seem quite exotic, but, as a matter of fact, the gender-based classifier system is widespread among Indo-European languages, and several languages incorporate it in their nominalization markers as Bora does. The Indo-Aryan language Marathi, for example, has argument nominalizations involving verbal forms traditionally labeled participle, which mark the gender of a denoted person. Compare the following examples.

(6-39) Marathi (Indo-Aryan; courtesy of Prashant Pardeshi)

a. [\(\ddot{d}\ddot{a}:\ddot{g}\ddot{i}n\ddot{e} \ s\ddot{\ddot{r}}\dddot{o}:l\ddot{e}:l\ddot{\ddot{\ddot{i}}}\)] q\(\ddot{a}:m\ddot{\ddot{c}}:\ddot{\ddot{i}}\) nok\(\ddot{r}i\)n \(\ddot{a}\)\(\ddot{h}\)e \\
\(j\ddot{e}:w\ddot{l} \ s\ddot{e}:l\ddot{a}: \text{pre.sent.-\text{m.s}} \ \text{we-gen-\text{m.s}} \ \text{female.servant} \ \text{be-\text{m.s}}\)

'The one (fem) who stole the jewels is our maid.'

b. [\(\ddot{d}\ddot{a}:\ddot{g}\ddot{i}n\ddot{e} \ s\ddot{\ddot{r}}\dddot{o}:l\ddot{e}:l\ddot{\ddot{\ddot{i}}}\)] b\(\ddot{a}\ddot{i} \ \text{m\dddot{a}:c}:\dddot{\ddot{c}}:\ddot{\ddot{i}}\) nok\(\ddot{r}i\)n \(\ddot{a}\)\(\ddot{h}\)e \\
\(j\ddot{e}:w\ddot{a}:l: \ s\ddot{e}:l\ddot{a}: \text{pre.sent.-\text{m.s}} \ \text{woman} \ \text{we-gen-\text{m.s}} \ \text{female.servant} \ \text{be-\text{m.s}}\)

'The woman who stole the jewels is our maid.'

(6-40)

a. [\(\ddot{d}\ddot{a}:\ddot{g}\ddot{i}n\ddot{e} \ s\ddot{\ddot{r}}\dddot{o}:l\ddot{e}:l\ddot{\ddot{\ddot{i}}}\)] \(\ddot{a}:\dddot{m}:\dddot{c}:\dddot{\dddot{\dddot{a}}}\) nok\(\ddot{r}a\)r \(\ddot{a}\)\(\ddot{h}\)e \\
\(j\ddot{e}:w\ddot{a}:l: \ s\ddot{e}:l\ddot{a}: \text{pre.sent.-\text{m.s}} \ \text{we-gen-\text{m.s}} \ \text{male.servant} \ \text{be-\text{m.s}}\)

'The one (masc) who stole the jewels is our male servant.'

b. [\(\ddot{d}\ddot{a}:\ddot{g}\ddot{i}n\ddot{e} \ s\ddot{\ddot{r}}\dddot{o}:l\ddot{e}:l\ddot{\ddot{\ddot{i}}}\)] m\(\dddot{\dddot{u}}\dddot{n}u\dddot{s} \ \text{m\dddot{a}:c}:\dddot{\dddot{c}}:\dddot{\dddot{a}}\) nok\(\ddot{r}a\)r \(\ddot{a}\)\(\ddot{h}\)e \\
\(j\ddot{e}:w\ddot{a}:l: \ s\ddot{e}:l\ddot{a}: \text{pre.sent.-\text{m.s}} \ \text{man} \ \text{we-gen-\text{m.s}} \ \text{male.servant} \ \text{be-\text{m.s}}\)

'The man who stole the jewels is our male servant.'
The Marathi argument nominalizations above also show that clauses/sentences do not underlie them. In Marathi, as in other similar Indo-Aryan languages, verbs agree in gender and number with an unmarked (or nominative) nominal argument. In the past tense, an A nominal is marked ergative and a P nominal is left unmarked, and in such a case a verb agrees with the P nominal, as shown below:

(6-41) a. mol.karṇi-ne dāgine tsor-le
   maid-erg jewels.mpl steal-pst.mpl
   ‘The maid stole the jewels.’

b. nokar-ne dāgine tsor-le
   male.servant-erg jewels.mpl steal-pst.mpl
   ‘The male servant stole the jewels.’

Comparison between these and the nominalized forms in (6-39) and (6-40) makes it clear that they do not underlie the latter, in which nominalized forms do not agree with the unmarked P argument.

German distinguishes three gender classes of masculine, feminine, and neuter, and, like ordinary nouns, grammatical argument nominalizations are distinguished according to these classes depending on what they denote. Observe:

(6-42) German
   a. Ich kenne den [der [Ø morgen kommt]].
      I know art.msc.acc msc.sub.nmlzr tomorrow comes
      ‘I know the one (msc) who comes tomorrow.’

   b. Ich kenne die [die [Ø morgen kommt]].
      I know art.fem.acc fem.sub.nmlzr tomorrow comes
      ‘I know the one (fem) who comes tomorrow.’

   c. Ich kenne das [das [Ø morgen kommt]].
      I know art.neut.acc neut.sub.nmlzr tomorrow comes
      ‘I know the one (neut) who comes tomorrow.’

The German nominalizers clearly combine information about the gender class and about the grammatical relation that the denoted entities are identified with. In fact, German has a double marking system, whereby nominalizations have the structure internal marking discussed above as well as the external marking by the articles indicating the gender class of the entity denoted by the nominalizations, as observed in (6-42) above.\(^{59}\) In some languages, external classificatory articles of the German type are the only clue to the nature of the entities denoted by grammatical argument nominalizations.

\(^{59}\) The articles marking these argument grammatical nominalizations, based on demonstrative pronouns, differ from those marking ordinary nouns, though there is a great deal of overlap in form and function between the two sets.
In Toba, all nominals in NP-use are marked by what Messineo (2003) calls nominal classifiers (clasificadores nominales), which encode configurational and deictic, as well as number and gender information regarding the denotation of the following nominal, as below.60

(6-43) Toba (see Messineo (2003: 145ff) for details)

a. \( na \) \( pioq \)
   \( DD \) \( dog \)
   PROXIMAL
   ‘this dog’

b. \( yi-wa \) \( pioq \)
   \( DD-PAUC \) \( dog \)
   HORIZONTAL
   EXTENDED
   ‘two or three dogs lying down’

c. \( a-so \) \( waaka \)
   FEM-\( DD \) \( cow \)
   DISTAL
   ‘a cow’

Just like the German articles mentioned above, these demonstrative determiners indicate the nature of the entity denoted by grammatical nominalizations, as seen below.

(6-44) Toba (courtesy of Cristina Messineo)

a. \( s-ac\text{e}k \) \( a-so \) \( [(ntonigifi) [\text{aw-\text{\text{-}}ot} \ (fikaj]\_\text{\text{-}}\text{nmlz}]_\text{\text{-}}\text{nmlz}]_\text{\text{-}}\text{NP} \)
   1A-eat \( FEM-\text{DD} \) \( (\text{tortilla}) \)
   2A-make \( \text{yesterday} \)
   DISTAL
   ‘I ate what you made yesterday.’ ‘I ate the tortilla that you made yesterday.’

b. \( [na-wa] [\text{\text{-}}chigoqchigm\mu yi \ (Espinillo)]_\text{\text{-}}\text{nmlz}]_\text{\text{-}}\text{nmlz}]_\text{\text{-}}\text{NP} \)
   \( DD-\text{PL} \)
   3.come.from \( DD \) \( (Espinillo) \)
   PROXIMAL
   \( \text{\text{-}}\text{ayge} \)
   \( da \)
   \( \text{Salta ko'ollaGa} \)
   \( 3.\text{go<PAUC>} \) \( DD \)
   \( \text{Salta PST} \)
   ‘Those who came from the Espinillo went to Salta.’

These demonstrative determiners (DDs) (and the demonstratives not discussed here) show two things. Syntactically, grammatical nominalizations are nominal and they are marked by either a demonstrative determiner (or a demonstrative) in

---

60. Cristina Messineo (p.c.) now would call these “demonstrative determiners” (DDs) and “demonstratives” (DEMs). I gloss the examples with this new terminology.
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

Their NP-use, just like any ordinary nouns. In other words, determiners function as NP-use markers for both ordinary nouns and grammatical nominalizations. Semantically, grammatical nominalizations, especially argument nominalizations, denote concrete entities, whose physical properties including number and gender are marked by DDs (or demonstratives), again just like any ordinary nouns. These facts would not be easily explained if grammatical nominalizations were clauses and sentences that do not denote substantives. Notice also that grammatical nominalizations are perfectly compatible with finite verb forms in Toba as well.

We conclude this section by pointing out that English grammatical argument nominalizations also classify their denotations in terms of the human/non-human distinction similar to the animate/inanimate distinction that Newar nominalizers mark. The case in point is the distinction between who(m) and which, the former marking a human denotation by an argument nominalization and the latter a non-human denotation. Observe:

(6-45) a. You may marry [[who [you like Ø] nmlz]NP
b. You may choose [[which [you want Ø] nmlz]NP

This again shows that what we consider to be grammatical nominalizations are denoting rather than predicating or asserting like clauses and sentences.

7. Nominal-based nominalizations

Perhaps the most innovative proposal made in Shibatani’s work on nominalization is to reanalyze the genitive or possessive construction as a nominal-based nominalization. There are several motivations for this radical departure from the traditional analysis. First of all, what forms like his and John’s denote are those things with which the person referred to is intimately connected, as in the case of ordinary metonymic expressions, such as things that are possessed permanently or temporarily or things to which the person is connected as an author or a theme (as in the case of the theme of a photo).

---

61. See Section 8.1 on the development of these NP-use markers as nominalizers.
62. Movima (a language isolate spoken in Bolivia; Haude 2006) and Musqueam (Suttles 2004) and related Salish languages have demonstrative systems similar to Toba.
63. We are dealing with the genitive of “possession” in this article, not other kinds of genitive, such as, for example, the Russian genitive used in lieu of the partitive case.
Secondly, the precise references of the “possessive” forms are determined by context, again as in the case of ordinary metonymic expressions and verbal-based grammatical nominalizations discussed above.

Observe:

(7-1)  A: Which car do you like?
       B: I prefer John’s over Bill’s.

(7-2)  A: Is this the book that Bill brought?
       B: No, that’s John’s. I saw Bill’s on the dining table.

By the Gricean Cooperative Principle, we would interpret John’s and Bill’s in (7-1) to be referring to the cars intimately connected with the referents of John and Bill, while in (7-2) John’s and Bill’s would be likely understood to be referring to the books belonging to the referents of John and Bill.\(^{64}\) The relevant forms above represent NP-use of N-based nominalizations. These, like V-based grammatical nominalizations studied above, also have a modification-use, as shown below.

\[
\begin{array}{c|c}
\text{Structure} & \text{Use/Function} \\
\hline
\text{NP-use/Referring function} & \text{[[[John]s}\_NMLZ]\_NP is on the dining table.} \\
& \text{[[His}\_NMLZ]\_NP is this.} \\
\text{[[John]s}\_NMLZ} & \text{[[his}\_NMLZ} \\
\text{[his]} & \text{Modification-use/Restrictive function} \\
\text{NMLZ} & \text{[[[John]s}\_NMLZ \text{book}\_NMLZ]\_NP is on the table.} \\
& \text{[[His}\_NMLZ \text{book}\_NMLZ]\_NP is this.}
\end{array}
\]

Figure 4. Two uses of N-based nominalization

The newly proposed nominalization analysis does away with the genitive case for possessive constructions altogether as well as the parts of speech of “possessive pronouns/pronombres posesivos” and “possessive adjectives/adjetivos posesivos” recognized in traditional grammar. The former (his, mine, etc.) are no more than instances of the NP-use and the latter (his, my, etc.) those of the modification-use

---

\(^{64}\) These nominal-based nominalizations in NP-use are described variously as “free genitives” (e.g. Stolz et al. 2008: 390), “headless adnominal” (e.g. Noonan 2008: 130), and “headless possessive” (Koptjevskaja-Tamm 1995; van der Voort 2009). See also Dryer (2008). The same arguments against such treatments can be raised as those against treating the NP-use of argument nominalizations as “headless relative clauses” in Section 5.4.
of N-based nominalizations. Traditional grammar makes the same mistake as those who recognize relative clauses apart from the modification-use of grammatical argument nominalizations.

Besides the logical consistency between the analysis of N-based nominalizations outlined above and that of V-based nominalizations, Shibatani (2017, 2018a) offers morphological evidence unifying these two types of grammatical nominalization. The NP-use of V-based nominalizations in Modern Japanese requires no-marking very much similar to the one-marking found in Modern English. Compare the forms of the two uses of V-based nominalizations in the Japanese forms and their English translations below.

(7-3) Japanese
   a. \[ ([suki \text{nmlz}]_{NP} \text{no to kekkon sinasai}.) \]
   \[ \text{like} \ \text{COP} \ \text{NPM with marriage do.IMP} \]
   ‘Marry one [who [you like]_{NMLZ}NP]
   b. \[ ([suki \text{nmlz} \text{hito}]_{NP} \text{to kekkon sinasai}.) \]
   \[ \text{like} \ \text{COP} \ \text{person with marriage do.IMP} \]
   ‘Marry [a person [who [you like]_{NMLZ}NP]

Shibatani traces the no-marker above to the Classical Japanese no that marks the NP-use of the N-based nominalization (or the genitive form), as seen in an example like the one below.

(7-4) Hitozuma to [wa ga\(^{65}\)]
   man’s.wife and I gen
   no hutatu omouni hanarekosi sode wa awaremaseru.
   NPM two think leave.behind sleeve TOP exceedingly.sad
   ‘As I think about both a man’s wife and mine, the sleeves left behind are exceedingly sad.’ (Yoshitadashū, 10th C)

The extension of the no-marking from the NP-use of N-based nominalization (aka the genitive/possessive), as in (7-4) above to that of the V-based nominalization, as in (7-3a) started in the early 17th century. Shibatani’s point is that this extension of the no-marking from one domain to another indicates that the two domains are recognized as a unified phenomenon.

Crosslinguistic investigations reveal a large number of cases where both N-based nominalizations and V-based nominalizations take the same morphological marking. There are two patterns of marking here. One is the Japanese

---

\(^{65}\) Classical Japanese had two genitive particles (or nominalizers for nouns), no and ga. The no particle that marks the NP-use of grammatical nominalization is related to the nominalizer/genitive no. In some other dialects (e.g. Kōchi, Toyama) ga is used as the marker of the NP-use of grammatical nominalizations.
pattern, where only the NP-use of both N-based and V-based nominalizations are marked the same way, and the other, perhaps more compelling pattern is where both N-based and V-based nominalizations involve identical markings in all usage patterns.

The Korean use of the particle *kes* is similar to the Japanese *no*-marking, where only NP-uses of N-based and V-based nominalizations are marked identically, as below.

(7-5) Korean

a. NP-use of V-based nominalization

\[
Na-nun [[[apeci-ka cwu-\text{-}n]_{\text{NMLZ}} kes]_{\text{NMLZ}}]_{\text{NP}} ul \text{ ilk-ko-iss-ta.}
\]

I-\text{TOP} father-NOM give-NMLZR-NPM-ACC read-GER-be-IND

‘I am reading what father gave (me).’

a’. NP-use of N-based nominalization

\[
[[[emeni-uy]_{\text{NMLZ}} kes]_{\text{NMLZ}}]_{\text{NP}} un ku chaek i-ta.
\]

mother-NMLZR-NPM-TOP that book COP-IND

‘Mother’s is that book.’

b. Modification-use of V-based nominalization

\[
[[apeci-ka cwu-n]_{\text{NMLZ}} chaek]_{\text{NP}}
\]

father-NOM give-NMLZR book

‘the book that father gave (me)’

b’. Modification-use of N-based nominalization

\[
[[emeni-uy]_{\text{NMLZ}} chaek]_{\text{NP}}
\]

mother-NMLZR book

‘mother’s book’

Notice that the nominalizers themselves are different for the V-based (-\text{n}) and N-based (-\text{uy}) nominalization, yet the results of these processes are treated alike in their NP-use, as seen in (7-5a) and (7-5a’).

Many languages of the world (e.g. a variety of Japanese dialects, Ryukyuan, Telugu and some other Dravidian languages) show a similar marking pattern for the NP-use of both N-based and V-based grammatical nominalizations (see Section 6.1 for the Telugu pattern). While the ultimate origins of the Japanese particle *no* and Korean *kes* cannot be known, many languages recruit as markers of NP-use of nominalizations a noun meaning “thing”, as closely documented in a variety of Ryukyuan languages by Shibatani and Shigeno (2013).\(^{66}\) The Kwa language

\(^{66}\) Many Korean scholars think that *kes* was originally a noun with the meaning of “thing”, but there is no evidence for it. The “thing” reading they associate with *kes* actually comes from the nominal denotation of the nominalizations they mark. Like Japanese, Middle Korean did not have the *kes* marking, yet those nominalizations without *kes* have exactly the same “thing” reading as their modern counterparts with *kes*. See Lee (1975).
Gã of Ghana uses *nɔ́*, deriving from a noun meaning “thing”, and *mɔ́*, which means “person” as a noun, as markers of NP-use of nominalizations—the former when a non-human is denoted and the latter for a human denotation. While in Gã the origins of these markers are transparent, Campbell (2017: Chapter 6) presents strong evidence that they are grammaticalized and do not mean “thing” or “person” when they occur with the NP-use of nominalizations. For example, *nɔ́* as a noun takes a definite article but *nɔ́* as a marker of NP-use never does, and *nɔ́*, meaning inanimate “thing” as a noun, can mark both human and animate referents in the NP-use of N-based nominalizations for a smaller group of speakers.67

(7-6) Gã (Kwa; Campbell 2017: 111, 538, 550, 557)

a. NP-use of V-based nominalization

\[
\begin{align*}
\text{[nɔ́ [ni śūm3ɔ́ɔ́ɔ́] jí ãnǐháó]} & \\
\text{NPM NMLZR 1SG=like.NEG COP laziness} & \\
\text{‘What I don’t like is laziness.’}
\end{align*}
\]

a’. NP-use of N-based nominalization

\[
\begin{align*}
\text{shí [\{Ellen\] nɔ́]=! ṭ} & [\{Pápá Téi nɔ́\]=! ṭ lē=!č \\\n\text{but Ellen NPM=top Papa Tei NPM=top 3SG.OBJ=top} & \\
\text{i=nyééé mà-lyá} & \\
\text{1SG-able.NEG 1SG.FUT-go} &
\end{align*}
\]

‘But as for Ellen’s and Papa Tei’s I couldn’t attend them.’

b. Modification-use of V-based nominalization

\[
\begin{align*}
\text{[ətàlè [ní àmè=šùm3]} & \\
\text{dress NMLZR 3PL=like} & \\
\text{‘the dress that they like’}
\end{align*}
\]

b’. \([\{Elma\] bît\]68 jí lē

\[
\begin{align*}
\text{Ellma child COP 3SG.OBJ} & \\
\text{‘She’s Elma’s child.’}
\end{align*}
\]

---

67. Notice that these markers of the NP-use of nominalizations are the result of *appropriations of lexical nouns* meaning “thing”, “person”, “matter”, “place” etc. as a grammatical marker, rather than resulting from their lexical use through a gradual grammaticalization process. This is so because the lexical use of these nouns continues in parallel with the grammatical use, as in many Ryukyuan languages, where constructions marked by *mun(u)* meaning “thing” lexically can be ambiguous (e.g. *wa: munu* ‘my thing/mine’). In the Taketomi Ryukyuan this ambiguity has been resolved by changing the grammatical version as *ũnu* (*ba: munu* ‘my thing’ *vs. ba: ũnu* ‘mine’). See Section 8 below and Shibatani & Shigeno (2013) on the use of *mun(u)* in verbal-based nominalizations.

68. Notice that Gã, as in many languages, does not have an overt nominalization marker for N-based nominalizations.
The recruiting of a noun meaning “thing” as a marker of NP-use of nominalizations is also seen in the Panoan language Kakataibo (Zariquiey 2011).

Notice at this juncture that we are dealing with two kinds of nominalization markers—69—one, a marker of NP-use glossed as npm, which, as in the examples above, marks only the NP-use of nominalizations, and the other, a nominalizer (or nominalizing morphology) glossed as nmlzr, which marks nominalizations and which appears wherever they are used, e.g. in both NP-use and modification-use of nominalizations; see the nominalizers -n for V-based nominalizations and -uy for N-based nominalizations in the Korean examples in (7-5) above. It is mistaken to identify NP-use markers (npm’s), such as the Japanese particle no and Korean kes (as well as Gã nɔ̃) above, as nominalizers (as in, e.g. Horie 2008; Frellesvig 2010; Yap, Grunow-Hårsta & Wrona 2011).70

Both Japanese no in the central dialects and Korean kes are a modern development and the earlier (e.g. Middle Japanese and Middle Korean) NP-use of nominalizations was not marked by them, but they are now both obligatory as NP-use markers in the modern languages. Likewise, the English NP-use marker one seen in the translation of (7-3a) above is a modern development. Grammatical nominalizations marked by wh-forms were freely used as NP-heads without the one-marking in Middle English, and it is still optional in Modern English in certain contexts, as in many of the examples in this paper.71

The difference between NP-use markers and nominalizers can be clearly seen by comparing the marking pattern of Korean kes in (7-5) above and that of the Mandarin de nominalizer below, where de occurs in both NP- and modification-use.

69. We use the term “nominalization marker” as a cover term for both markers of nominalizations in NP-use, like Korean kes and Gã nɔ̃, and nominalizers, like Korean -n and -uy and Mandarin de, which occur in both contexts of NP- and modification-use.

70. Whether or not a morphology marking nominalization is an NP-use marker or a nominalizer cannot be determined without a systematic investigation of usage patterns of nominalization. It is, therefore, dangerous to rely on others’ descriptions of nominalization markers, as done in many research papers, since most past works on this topic have not done the necessary systematic investigation called for.

71. The Middle English Dictionary compiled by the University of Michigan (https://quod.lib.umich.edu/m/med/) contain forms such as And [which falleb on Fat furste flur] schal beo Quene and [Who aske this] Leseth his asking trewely, both of which would require marking by one in Modern English. Sinhala is another language that has recruited the numeral eka ‘one’ as a marker of the NP-use of both event and argument nominalizations, where the marking is obligatory.
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

115

(7-7) Mandarin Chinese

a. NP-use of V-based nominalization

Nie mei you \([\{wō xihuan O\} = de\]_{NMLZ]\)NP
you not have I like =NMLZR
‘You don’t have what I like.’

b. Modification-use of V-based nominalization

[wō xihuan O] = de\]_{NMLZ} yi fu\]_{NP}
I like =NMLZR clothes
‘the clothes that I like’

While the marking of both N-based and V-based nominalizations by the same nominalizing morphology has been noticed by many, including Matisoff (1972), there has been no answer as to why a V-based nominalization and a “possessive construction” are marked the same way.\(^{72}\) As for Mandarin Chinese, Li & Thompson (1989) recognize two different *de*, one for nominalizing verbs (p. 575ff) and the other termed “associative” *de* (p. 113ff) for N-based expressions, as if we are dealing with two different particles that are accidentally similar in form.\(^{73}\) Such a treatment cannot explain why a similar “accident” happens in so many languages across the globe. Sposato (2012), in his description of relative clauses of the Miao language Xong, opts for Li & Thompson’s term in describing one type of V-based argument nominalization and N-based nominalizations, both marked by what appear to be interchangeable markers *naond* and *nangd*, leaving unanswered the question why relative clauses and possessive constructions are marked the same.\(^{74}\)

\(^{72}\) DeLancey (1986: 1), maintaining a narrow, verb-centered view of nominalization, finds it “odd that a dependent noun [of possessive constructions] would be marked as nominalized”.

\(^{73}\) After completing this manuscript, I learned that Zhu (1982) had offered an analysis highly similar to ours, recognizing N-based nominalizations along with V-based ones in terms of the Mandarin particle "de" under discussion.

\(^{74}\) It is probable that the Xong pattern results from contact with Chinese. Sposato recognizes other functions these markers play, such as marking adverbs and functioning as a sentence final emphatic marker. The development of nominalizers into these functions is not at all rare (see Yap & Grunow-Härsta (2010)).
(7-8) Xong (Miao-Yao (Homong-Mien); southern China; Sposato 2012)

a. \[ Wud jangs nangd\]?\^\(3\) nis \ ndut-lid \ ndut-ghueax.  
\[
3\text{sg} \text{ plant } \text{assoc} \quad \text{cop tree-plum tree-peach}
\]
‘What he planted were plum trees and peach trees.’

b. \[ Wel hauk naond \] jud \ jix raun.  
\[
1\text{sg} \text{ drink } \text{assoc} \quad \text{alcohol } \text{neg} \text{ good}
\]
‘The alcohol that I’m drinking is no good.’

c. \[ Ob-naind nis [wel naond]. \]
\[
\text{nom-this} \text{ cop} \quad 1\text{sg} \text{ assoc}
\]
‘This is mine.’ (NOM=nominalizing prefix or general nominal prefix)

d. \[ dab-guoud naond \] zhoux.mioux  
\[
\text{AN-dog} \quad \text{assoc} \quad \text{ear}
\]
‘the dog’s ear’ (AN=animal prefix)

A wide range of Tibeto-Burman languages (e.g. Lahu, Burmese) show a pattern similar to the Chinese (Sinitic) and Xong (Miao-Yao) pattern above, marking both V-based argument nominalizations and N-based nominalizations the same way. While the data clearly cry out for uniform treatment of these marking patterns, the nominalization solution was unavailable to those who have puzzled over the relevant data because of the past beliefs that nominalization applies only to verbs or non-nominal elements. We have shown in the beginning of this paper that these beliefs are in fact groundless; and what we see in Chinese and Xong above (as well as Matisoff’s earlier observations on Lahu) confirm this.

Turning to languages outside the Sino-Tibetan sphere, it is not easy to find comparable cases among languages of the Americas largely because most descriptive grammars fail to record the NP-use of N-based nominalizations comparable in form to mine and the child’s, as in Mine/ The child’s is this one. However, at least two languages are found that show the use of the same marking pattern for both V-based and N-based nominalizations. Also, there are a fair number of languages that treat V-based and N-based nominalizations similarly in terms of classifier marking, as discussed below.

The first, a more straightforward, case found in the Mosetenan language Mosétén in Bolivia concerns the nominalizers -tyi’ (M) and -si’ (F), which classify their denotations according to their gender. While Sakel (2004) treats them as linkers, she

75. This and the example in (7-8c) show that the term/gloss “associative” is quite infelicitous if what is referred to by this term were to “indicate[s] that two noun phrases [connected by it] are ‘associated’ or ‘connected’ in some way” (Li & Thompson 1989: 113), because these examples do not involve two noun phrases connected by nangd or naond. On the other hand, if the term is understood to be referring to a nominalizer marking structures that denote entities metonymically associated with them, then it makes a good sense. But this is not what is intended by Li & Thompson’s use of the term.
clearly recognizes their nominalization function, citing forms like the ones below representing both V-based lexical and grammatical nominalizations. Also observed below are N-based grammatical nominalizations utilizing the same nominalizers.

(7-9) Mosetén (Sakel 2004: 97, 106, 107, 111)

a. V-based nominalization
i. \( mi' \ jae'ma [ji'-wë-ti-ksi-tyi'] \)
\[ 3m.sg \ dm \ ca-move-dt-3p.0.m.s-nmlzr \]
\( 'the one who made them come' \)
ii. \( Mi' \ soni' [yo-dye'-tyi'] \)
\[ 3m.sg \ man \ r-nmlzr-nmlzr \ m-1n live-\text{V}.m.s \ kill-\text{V}.3m.o \]
\( jii-ri-ty iitsiki. \)
\( jaiir-ty itsiki. \)
\( 'The man who lives in here killed a jaguar' \)

b. N-based nominalization
i. \( phen-mi' [jae'ma iits [Kose-si' jike] \)
\[ woman-3m.sg \ dm \ de.m \ jose-nmlzr \ ps \]
\( \text{Tureno-win} \)…\)
\( \text{Tureno-c} \)
\( 'his wife, the late Jose Tureno's' \)
ii. \( mo' [aka' jaiichbae'-si] Martin-si' \)
\[ 3f.sg \ house red-nmlzr \ martin-nmlzr \]
\( 'Martin's red house' \)

While Sakel (2004: 106) translates (7-9b.i) above as “the wife of the late Jose Tureno” in a modification pattern, a literal interpretation seems to be the one suggested for this form, where the possessor form is understood to be in an NP-use in apposition with the preceding noun phrase “his wife”. This seems correct in view of the occurrence of the demonstratives in (7-9b.i). As seen in (7-9a.ii) and (7-9b.ii), modifying nominalizations are not marked by demonstratives.

The other case marking N-based nominalization by the same nominalizer that marks V-based nominalization concerns the marking by the Tapiete nominalizer -wa, seen earlier (5-9). The situation is not as straightforward as the Mosetén case above. The NP-use of possessive forms involves the extra morpheme \( a'\text{ampo} \), which González (2005: 243) simply glosses as “POSS”, as below.

(7-10) Tapiete (González 2005: 114, 243)
\( mbo'u she sh-a'\text{ampa-wa} \)
\( 'Pass me mine' \)

76. See the Newar forms in (7–91) below that also show a doubling of nominalization markers.
Turning to languages outside the Americas, first observe the comparable marking pattern in the Niger-Congo language Yoruba in West Africa, where the nominalizer for V-based nominalizations has high tone, whereas the counterpart for N-based ones has mid tone.

(7-11) Yoruba (Ajiboye 2005)

a. V-based nominalization

\[
\text{Mo ri eyi [tí Kúnlé ni]} \\
\text{I see this NMLZR Kunle own} \\
\text{‘I saw the one that Kunle owns.} \\
\text{Cf. [ère [tí Kúnlé ni]]} \quad (90) \\
\text{statue NMLZR Kunle own} \\
\text{‘the statue that Kunle owns’}
\]

b. N-based nominalization

\[
\text{Mo ri [ti Kúnlé]} \quad (107) \\
\text{1sg see NMLZR Kunle} \\
\text{‘I saw Kunle’s.’} \\
\text{Cf. [ère [ti Kúnlé]]} \\
\text{statue NMLZR Kunle} \\
\text{‘Kunle’s statue’}
\]

The Cushitic language Kambaata marks nominalization by a final accent in both V-based and N-based nominalizations.

(7-12) Kambaata (Cushitic; Ethiopia; Treis 2008)

a. V-based nominalization

i. \[
[[\text{dagujj-ó]} \quad \text{adab-áa}] \\
\text{run-3M.PVO.NMLZR boy-M.ACC} \\
\text{‘the boy who ran’}
\]

ii. \[
[[\text{xuujj-o-sé]} \quad \text{adab-áa}] \\
\text{see-3M.PVO-3E.OBJ.NMLZR boy-M.ACC} \\
\text{‘the boy who saw her’}
\]

b. N-based nominalization

i. \[
[\text{ann-í}] \quad \text{hiz-óo} \\
[\text{N-M.GEN.NMLZR} \text{ N-M.ACC}] \\
\text{‘father’s brother’}
\]

ii. \[
[\text{ann-i-sé}] \quad \text{hiz-óo} \\
[\text{N-M.GEN-poss.NMLZR} \text{ N-M.ACC}] \\
\text{‘her father’s brother’}
\]
iii. \([\text{ann-a kk-a}-\text{s}\text{é}]\) \(\text{hiz-\text{o}}\)  
\([\text{N-PL-F.GEN-POSS.NMLZR}]\ \text{N-M.ACC}\)  
‘her fathers’ brother’

The Austronesian language Lamaholot in eastern Indonesia makes use of another kind of suprasegmental feature to mark nominalization of both verbs and nouns, namely nasality, which is likely connected to the Proto-Austronesian genitive determiner *ni. Observe:

(7-13) Lamaholot (Austronesian; Nagaya 2011: 194, 200)

a. V-based nominalization  
\(\text{go hope} [\text{meʔ\text{å}] / topi meʔ\text{å}]\).  
1SG buy red.NMLZR / hat red.NMLZR  
‘I bought the red one/red hat.’

b. N-based nominalization  
\(\text{go gute Hugo nəʔ\text{ē}}\) \((\text{hepe}).\)  
1SG take Hugo 3SG.NMLZR knife  
‘I will take Hugo’s (knife) [lit. ‘Hugo his (knife)’].’

The Indo-Aryan language Nepali, likely influenced heavily by surrounding Tibeto-Burman languages, marks both V-based argument nominalizations and event nominalizations the same way as N-based nominalizations, as shown below.

(7-14) Nepali (courtesy of Madhav Pokharel)\(^\text{77}\)

a. V-based event nominalization  
\(\[[\text{u Dhilo aa-}\text{e}=\text{ko}]\] \(\text{durbhaagya bha-\text{yo}.}\)  
s/he late come-PF-NMZLR unfortunate BE-PFV  
‘That s/he came late was unfortunate.’

b. V-based argument nominalization  
\(\[[\text{Madhav le ma laai di-}\text{e}=\text{ko}]\] \(\text{ma paDh-dai} \text{chu}\)  
Madhav ERG I DAT gv-PF-NMLZR I read-PROG am  
‘I am reading what Madhav gave me.’

Cf. \(\[[\text{Madhav le ma laai di-}\text{e}=\text{ko}]\] \(\text{kitab}]\)  
Madhav ERG I DAT gv-PF-NMLZR book  
‘the book that Madhav gave me’

c. N-based nominalization  
\(\text{[Madhav}=\text{ko}] \text{ ma paDh-dai} \text{chu}\)  
Madhav=NMLZR I read-PROG am  
‘I am reading Madhav’s.’

\(^{77}\) To appreciate a systematic different between nominalizers (NMLZRs) and NP-use markers (NPMs), compare these Nepali examples with a NMLZR and the Telugu examples in (8–1) with an NPM.
A comparable pattern is seen in Modern Hebrew, in which she marks a similar range of nominalizations as in Nepali (also see Shibatani & bin Makashen (2009) for another Semitic language Soqotri).

(7-15) Modern Hebrew (courtesy of Ana-Marie Hartenstein)

a. V-based nominalization
   i. Ani yodaat [she] [ata lo betem rofe]  
      ‘I know you a real doctor’
   ii. zo [she] [Yoav raa etmol] xi xauver shel 
      ‘The one Yoav saw yesterday is a friend of mine.’

b. N-based nominalization
   [Ha-kova [she-l Moshe]] shachor, aval [she-l 
   art-hat nmlzr-dat? Moshe black but nmlzr-dat? 
   Yakov] chum. 
   ‘Moshe’s hat is black but Yakov’s is brown.’

Next, those languages that use classifiers as nominalizers may mark both V-based and N-based nominalizations by classifiers, as in Cantonese below.

(7-16) Cantonese (Matthews & Yip 1994: 108, 111, 112)

a. V-based nominalizations marked by classifiers
   i. [Ngó-deih hái Faatgwok sihk] di yéh géi hóu-sihk ga. 
      we in France eat clf food quite good-eat prt 
      ‘The food we ate in France was pretty good.’
   ii. [[Gaau léih tàahn kàhm gò] go? 
      teach you play piano that clf 
      ‘The one who teaches you piano?’

b. N-based nominalizations marked by classifiers
   i. [léih (gó) di] pàhngyáuh 
      you (that) clf friend 
      ‘those friends of yours’
   ii. [léih (gó)] pàhngyáuh 
      you clf friend 
      ‘your friend’
iii. Lī dēoi hai [ngóh dī phàngyáuh], [léih gó dī] these (lit. This pile) COP I CLF friend you that CLF hōeng gōpihn. LOC there (courtesy of Haowen Jiang)

‘These are my friends, and yours are over there.’

Similar use of classifiers is also seen among Amazonian languages, as shown by the Tucano language Barasano (also known as Pãnerã) in Colombia.

(7-17) Barasano (Tucano; Jones & Jones 1991: 61,150)

a. V-based nominalization

[hũ ʉ hammock] ŏ kāhi-ri-ku ābo-a-ha yu
hammock there hang-NMLZR-CLF want-PRS-3 1SG
‘I want the hammock that is hanging there.’

b. N-based nominalization

[hũ ʉ hammock] ţ-ya-g ʉ
hammock 3MASC.SG-NMLZR-CLF
‘his hammock’

Barasano has different nominalizers for V-based (-ri) and N-based (-ya), as do many other languages. However, the results of the nominalizations are treated alike, as indicated by use above of the same classifier marking ku/ɡu, which is for a long hammock. Both V-based and N-based forms have NP-use such that (7-17a), without the head noun hũ, would mean “I want one (hammock-thing) hanging there” and (7-17b) “his (hammock-thing)”.

In the related language Tuyuca, classifier marking is optional for the N-based nominalization in modification-use, while it is obligatory in the NP-use, as seen below.78

(7-18) Tuyuca (Tucano; courtesy of Janet Barnes)

a. V-based nominalization

[nikā [bako-ā-ri-gi]]
leg to.have.been.bitten-RECENT-SG.NMLZR-CLF
<cylindrical.shape,long.and.solid>
‘the leg that was bitten’

b. N-based nominalization

i. [yiistikə-yə-ró]NMLZRNP
my father-NMLZR-CLF:2D.FLEXIBLE
‘my father’s (as in “They are my father’s/My father’s are those.”)

78. In Bora, only nominal-based nominalizations appear to be marked by classifiers only in their NP-use.
This is the pattern that we find in the Tibeto-Burman language Newar, which has classifier-based nominalizers. Observe the data below, where an N-based form has its own nominalizer (-yā), but it further takes the nominalizer marking V-based nominalization (-mha), indicating that N-based nominalizations are treated like V-based nominalizations.

(7-19) Newar (courtesy of Kazuyuki Kiryu)

a. [[(ana dan-ā cwā=mha) macā] [rām=yā]=mha kha:]
   there stand-cm exist.nd=NMLZR=child Ram=NMLZR=NMLZR cop
   ‘The child standing over there is Ram’s.’

b. [[(ana dan-ā cwā=mha) rām=yā(mha)] macā kha:]
   there stand-cm exist.nd=NMLZR=child Ram=NMLZR=NMLZR cop
   ‘The one standing there is Ram’s child.’

Finally, Bantu noun-class marking, which also has a nominalizing function, marks both V-based and N-based nominalizations, as shown by the Chichewa examples below.

(7-20) Chichewa (Mchombo 2004 and p.c.)

a. V-based nominalization in NP-use
   7-NMLZR I-pst-buy 7-pst-be 7-expensive
   ‘What I bought was expensive.’

a’. V-based nominalization in modification-use
   7-hat 7-NMLZR I-pst-buy 7-pst-be 7-expensive
   ‘The hat that I bought was expensive.’

b. N-based nominalization in NP-use
   [[(ch-ângâ)nmlz]NP chí-ma-sangalâts-â a-lenje.
   7-my 7-HAB-please-fv 2-hunters
   ‘Mine pleases hunters.’

b’. N-based nominalization in modification-use
   7-hat 7-my 7-HAB-please-fv 2-hunters
   ‘My hat pleases hunters.’

All in all, there is ample evidence that languages around the globe also nominalize nouns and noun phrases. Our reanalysis of the so-called genitive case as an N-based nominalizer not only captures the parallel patterns exhibited by V-based and N-based nominalizations we have examined above, but also offers a new insight into the meaning relationship between the “possessor” and the “possessum”. Those various “possessive” relations (John’s head, John’s father, John’s letter, John’s hospital, etc.) are attributed to the metonymic inferences associated with a nominalization process creating the so-called possessive form, John’s. Modification in possessive constructions means restricting the denotation of the head noun (e.g. book) with that of the N-based nominalization (e.g. John’s), with the latter denoting things with which John is associated and specifying a subset of the former. The denotation of an entire possessive construction (e.g. John’s book) is an intersection of the denotation of a head noun and that of an N-based nominalization, in the same way as the so-called restrictive relative clause restricts the denotation of the head noun (see Section 5.5).

Traditional grammars treat the genitive case/possessive form as an inflection similar to a grammatical case such as nominative and accusative. The genitive, however, differs from grammatical case in that it either modifies another noun (his/John’s car) or stands in NP positions where case inflected forms cannot freely stand (His/John’s/*Him is here; I saw his/John’s/*he), showing that it is different from ordinary case forms, the syntactic positions of which are fixed – a nominative form in subject position, an accusative form in object position. The genitive form, in contrast, is not bound to any particular syntactic position, and even to the modifier position in possessive constructions, as the foregoing discussions amply demonstrate. More importantly, the denotation of the nominal in different case forms remains constant, with case inflections adding only grammatical meanings. For example, the nominative I marks a subject function, and the accusative me an object function, while maintaining their denotation constant, namely the speaker. This is not so with the genitive form; mine does not denote the speaker but rather something metonymically related to the speaker. The same applies to languages using particles to mark case, as in Japanese forms boku ga (I NOM) and boku o (I ACC) vis-à-vis boku no (o mite) ‘(look at) mine’.

The nominalization analysis proposed above treats the genitive case form as derivational, as nominal-based nominalization, similar to the derivations of pig > piglet, village > villager, which derive new nouns with new denotations, unlike the inflections such as hel/him, pig/pigs, which do not change the denotations,

---

80. See next section on forms such as my/mine that are associated with specific functions.

but which differ only in grammatical meanings (subject/object, singular/plural). The same applies to Japanese, Korean and other languages that make use of particles. While Japanese particles \textit{ga} (nominative), \textit{o} (accusative), \textit{ni} (dative) are case particles, the so-called genitive case particle \textit{no} is not. Like the English possessive clitic ’s, it is derivational; \textit{John no} does not denote/refer an individual named John but things that are metonymically related to that person.

8. Structure, use, and form

One of the most interesting facts to observe in crosslinguistic research is the way languages respond to unity and divergence of function in terms of linguistic forms. The functional unity underlying nominalization phenomena is the creation of nominal structures denoting entity concepts. The divergences stem from several factors. One is the difference in input, i.e. verbal-based or nominal-based. The outputs of the nominalization process are of several types. We have distinguished between lexical and grammatical nominalizations, the former whose form-meaning relationships are entrenched and whose grammatical properties are highly similar to ordinary nouns, and the latter non-lexical grammatical structures created for the nonce. Of the grammatical nominalizations, there are event nominalizations and argument nominalizations. These nominalizations play different functions depending on their use, in particular whether they head a noun phrase, where they play referential function, or they modify a head noun, either restricting the denotation of the head noun (so-called restrictive relative clauses) or identifying the denotation of the head noun (so-called non-restrictive or appositive relative clauses). Careful observations of the data discussed above indicate that languages respond to these functional similarities and divergences in different ways. An interesting question to be raised is whether general crosslinguistic patterns emerge on the basis of which we might be able to make predictions about change in form over time.

As for the distinction between lexical and grammatical nominalizations, languages often make a clear formal distinction between the two. Yet, we have seen that a fair number of languages do formally express the functional unity between the two by marking both types in a morphologically uniform way. Indeed, in some cases the form is ambiguous allowing either a lexical or grammatical interpretation. For example, the Central Alaskan Yup’ik forms in (2-2) and the Mayrinax Atayal form in (3-12) can be interpreted either lexically (in the sense of the word, e.g. “child” or “singer”) or grammatically (in the sense of “one who is small” or “one who sings”). In situations like this, it is likely that grammatical nominalizations give rise to lexical nominalizations, where a form denoting an entity in an analytic manner has been applied to an entity whose meaning is not entirely
compositional, as in the case of designating a person who sings routinely or whose singing constitutes a professional activity. The opposite direction of development, where the marking of lexical nominalizations has been extended to grammatical nominalizations needs to be documented.82

Turning to the distinctions between nominal-based and verbal-based nominalizations and between event and argument nominalizations, many languages of the world make clear formal distinctions in them. But, again, a fair number of languages from different parts of the globe formally express functional unity by morphologically marking them in a similar way. We recognize two patterns of formal identity across these types of nominalization. One pattern expresses the fundamental functional unity that binds all these types of nominalization (i.e. a marker indicating that they are all nominalizations), using the same nominalizing morphology for all of them, as in Nepali (see (7-14)) and Modern Hebrew (7-15), among others. The other, perhaps more widespread pattern responds to the commonality in their use function by marking the same all these types of nominalization (only) when they are in NP-use. This can be seen most clearly in the Telugu forms below, where the particle di marks the shared referential function of nominalizations in NP-use in a uniform manner, as below. Some other Dravidian languages, Korean (7-5), Gã (7-6), and a variety of Ryukyuan (Shibatani & Shigeno 2013) and Japanese dialects (Shibatani 2017) show this marking pattern.

(8-1) Telugu (courtesy of K. V. Subbarao)

a. N-based nominalization in NP-use

idi   *naa-di*

this my-NPM

‘This is mine.’

Cf. *naa pustakam*

my book

‘my book’

b. V-based event nominalization in NP-use

[vāḷḷu   *vacc-in-a*-di   naaku telusu.]

they.NOM come-PST-NMLZR-NPM to me known

‘I know that they came.’

Cf. [vāḷḷu   *vacc-in-a*   sangati]

they.NOM come-PST-NMLZR news

‘the news that they came’

82. The English grammatical nominalizations of the form [singing of the national anthem] appears to be an innovation based on the lexical counterparts involved in [singing of the national anthem] and the like. The parallel grammatical nominalizations do not obtain in other Germanic languages like German and Swedish.
Many languages make a clear formal distinction between N-based and V-based nominalization, and between V-based event nominalization and V-based argument nominalization. English and many others have special forms (known as the genitive case or possessive form) for N-based nominalization distinct from those for V-based nominalizations (traditionally referred to as gerundive, participial, infinitive, or adnominal). A comparison of Portuguese and Spanish pronoun-based nominalizations reveals how different languages respond differently to the functional demand at two different levels. Portuguese, like Italian, maintains the formal uniformity in favor of formally expressing the functional unity underlying the pronoun-based nominalizations—that the relevant forms are of the same substance regardless of their use; NP-use: Olhe para o [meu] ‘Look at mine’; Modification-use: Olhe para [meu livro] ‘Look at my book’.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Use/Function</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NP-use/Referential</td>
<td>[meu], [teu], [seu], [nosso]</td>
</tr>
<tr>
<td>[meu], [teu], [seu], [nosso]</td>
<td>Modification-use/Restrictive</td>
<td>[meu], [teu], [seu], [nosso]</td>
</tr>
</tbody>
</table>

Figure 5. Portuguese pronoun-based nominalizations for first- (meu), second- (teu), third-person (seu) singular masculine forms and first-person plural masculine form (nosso)

On the other hand, Spanish, like French (and English), differentiates forms according to their use/function so as to express formally the difference in the usage and function; MP-use: Mira el [mío] “Look at mine”; Modification-use: Mira [mi libro] “Look at my book” (see next page).

When a language develops usage-based specialized forms, they are likely to acquire new grammatical properties, just as the Spanish NP-use form mío or its English counterpart mine cannot modify a noun. Recall that Northern Paiute
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

Event nominalizations marked by -na allows an adverbial function, including chaining constructions (see (5-58)). The forms performing this chaining function denoting/referring to simultaneous events are identical with those functioning as verb complements, and so the -na forms bearing the adverbial and V-complement NP function look alike. Northern Paiute has developed a specialized form marked by -si for events holding a sequential temporal relation (see (5-62)). Such a specialized form would not be able to function as an NP-head functioning as a syntactic argument.

A similar but a subtler pattern is observed about the specialized adverbial nominalizations marked by the conjunctive particle -te/-de in Japanese (see (5-57)). While they show a measure of syntactic nominality in patterning like ordinary nouns in the formation of noun-modifying forms, they are unlike ordinary nouns or straightforward event nominalizations in that they cannot head an argument NP. Compare:

(8-2) Japanese
a. Noun and nominalization-based adverbial in modification-use
   i. [[[eiga] no] kaeri-miti
      movie NMLZR return-road
      (lit.) 'a return-road of a movie'/ 'a return road taken after a movie'
   ii. [[[eiga o mi=te] no] kaeri-miti
       movie ACC watch=CON NMLZR return-road
       (lit.) 'a return road of having watched a movie'/ 'a return road taken after having watched a movie'

b. Noun and nominalization-based adverbial in NP-use
   i. [[[Eiga]NP wa omosiroi.
      movie top fun
      'A movie is fun.'

<table>
<thead>
<tr>
<th>Structure</th>
<th>Use/Function</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NP-use/Referential</td>
<td>[mío], [tuyo],</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[suyo], [nuestro]</td>
</tr>
<tr>
<td></td>
<td>Modification-use/Restrictive</td>
<td>[mi], [tu],</td>
</tr>
<tr>
<td>[mi], [tu],</td>
<td></td>
<td>[su], [nuestro]</td>
</tr>
</tbody>
</table>

Figure 6. Spanish pronoun-based nominalizations for first- (mi), second- (tu), third-person (su) singular masculine forms and first-person plural masculine form (nuestro)
Caution must be taken in that even those forms that function as NP-heads may not behave alike since NPs themselves can function differently. The Portuguese pronoun-based nominalizations discussed above, such as *meu* 'my/mine (masc)' and *minha* 'my/mine (fem)', display different syntactic properties depending on different uses of NPs they head. When they head an NP functioning as a syntactic argument, they require a definite article, but when they head an NP functioning as a nominal predicate, article marking is optional. Compare:

(8-3) Portuguese

\[ a. \quad [\text{O meu}]_{\text{NP}} \quad \text{é aquele carro.} \]

the.msc 1.sg.msc.nmlzr is that car

'Mine is that car.'

\[ b. \quad \text{Aquele carro é } [\text{(o) meu}]_{\text{NP}} \quad \text{car}\]

that car is (the.msc) 1.sg.msc.nmlzr

'That car is mine.'

8.1 From NP-use markers to nominalizers

As a way of showing the importance of N-based nominalizations, the remainder of this subsection examines crosslinguistic marking patterns in grammatical nominalizations. Shibatani’s study of Ryukyuan and Japanese dialects (Shibatani & Shigeno 2013) shows that forms in NP-use tend to be more complex than the ones seen in modification-use, the former with an explicit marker for the referential function that the forms in NP-use bear (cf. Spanish *mi* 'my' vs. *mío* 'mine' and their English equivalents). Over time, however, the complex forms in NP-use spread to the modification context, leveling the formal difference between the two uses. Shibatani and Shigeno (2013) also show that once a language achieves uniformity in formal marking across different uses of the same type of nominalization, it begins to differentiate forms according to the difference in use, by adding an extra marker to the forms in NP-use. This cyclic development seems to be one way for a language to negotiate with the opposing forces for formal uniformity (revealing the underlying unity) and divergence (expressing the difference in usage/function).

Shibatani and Shigeno’s work demonstrates that the NP-use of N-based nominalization is the site where innovations start out, which then spread to the modification-use of N-based nominalizations, as well as to the NP-use of V-based nominalization, and then to their modification-use, as in the manner below.
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

The (a) pattern of development is clearly seen in a recent innovation in Úyama Okinawan on the main island of Okinawa in the Ryukyuan archipelago. There are two NP-use markers, *si* and *mun(u)*, seen in both N-based and V-based nominalizations that have a wide geographical distribution in the archipelago. The fact that *si* and its phonological variants occur only in peripheral areas such as southern Amami, Yoron, and Kumejima in the north and Ishigaki and Taketomi in the south indicate that it is older than *mun(u)*, whose occurrence with N-based nominalizations is seen in both central Ryukyuan islands and peripheral areas. The fact that while the etymology of *si* is unknown, *mun(u)* is transparently connected to the noun *mun(u)* ‘thing’ corroborates the observation that the *mun(u)* marking is a newer development.83 The marking patterns exhibited by *si* and the more widespread use of *mun(u)* for N-based nominalizations are shown below.

(8-4) Yoron Ryukyuan (Yoron Island, Kagoshima prefecture, Japan)
   a. NP-use of N-based nominalization
      \[\text{sinsee nu}]=\text{si} \quad \text{ja are}\]
      teacher NMLZR=NPM TOP that
      ‘The teacher’s is that.’
   b. Modification-use of N-based nominalization
      \[\text{sinsee nu} \quad \text{hasa}\]
      teacher NMLZR umbrella
      ‘teacher’s umbrella’

(8-5) Agena Okinawan (Okinawa Island, Okinawa prefecture, Japan)
   a. NP-use of N-based nominalization
      \[\text{ure: [sinsi: ga]} \quad \text{mun=do:}.\]
      that teacher NMLZR NPM=COP
      ‘That is the teacher’s.’

83. The use of the marker -*si* is attested in *Omorosōshi*, a collection of Ryukyuan songs compiled between 1531 and 1623, while the beginning use of *mun(u)* is unclear, largely due to the difficulty in determining whether *mun(u)* is functioning as a noun with the meaning of “thing” or as an NP-use marker.
b. Modification-use of N-based nominalization

\[ \text{ure: } \left[ [\text{sinsi: ga} \quad \text{kucu}] = \text{do}. \right. \]

that teacher NMLZR shoe = COP

‘Those are the teacher’s shoes.’

The following data show that the (a) pattern of development illustrated in Figure 7 is seen in the Ūyama dialect of Okinawan, where the mun-marking is spreading to the modification context, as seen in (8-6c) below, where it has replaced the N nominalizer.

(8-6) Ūyama Okinawan (Okinawa Island)

a. \( \text{Unu kucu ja } [\text{waN}] \text{ mun=do}. \)

that shoe TOP I.NMLZR NPM=COP

‘Those shoes are mine.’

b. \( [[\text{waN}] \text{ kucu}] \)

I.NMLZR shoe

‘my shoes’

c. \( [[\text{wa mun}] \text{ kucu}] \)

I NMLZR shoe

‘my shoes’

The replacement of the N nominalizer by mun has not yet developed to its full extent in that only first and second person pronominals can be marked by mun in the modification context. The fact that the neighboring dialect on Tsuken Island does not show this pattern of spread of the NP-use marker mun to the modification context shows that the above is a true innovation in Ūyama Okinawan.

A similar development illustrating the (a) pattern in Figure 7 is seen in Creek, which involves the noun root etymologically meaning ‘thing’ in the formation of the N-based nominalizations in NP-use, as below.

(8-7) Creek (Martin 2011: 144)

i. \( \text{ca-ná:ki ‘mine’} \)

1SG-THING

ii. \( \text{ci-ná:ki ‘yours’} \)

2-THING

iii. \( \text{i – ná:ki ‘his/hers/their’s’} \)

3-THING

iv. \( \text{po-ná:ki ‘ours’} \)

1PL-THING

v. \( \text{ca-ná:ki-t ò-s} \)

1S.PAT-THING-T be.FGR-IND

‘It’s mine.’
These forms contrast with those in the basic modification pattern, which marks the modified noun by personal prefixes directly, as below.

(8-8) a. relational noun
   i. ca-hácko ‘my ear’
   ii. ca-cá:ta ‘my blood’

b. non-relational noun
   i. am-ifa ‘my dog’
   ii. am-mí:kkó ‘my chief’

However, Creek allows what Martin (2011: 137–138) calls periphrastic possession using the NP-use forms in (8-7). Observe:

(8-9) a. tol-sakká:ka [ca-ná:ki] eye.glasses 1S.PAT-THING ‘my glasses’

b. ifá [ca-ná:ki-n] i:st o:w-éy-s dog 1S.PAT-THING-N take.SG.FGR-T be.1SG-1S.AG-IND ‘I’m holding my dog.’

According to Martin (2011: 138), some speakers prefer these periphrastic possessive forms for some non-relational nouns over the prefixed forms, as in (8-8), indicating that a shift from prefixed forms to periphrastic forms may be happening, as a way of leveling the formal difference between the NP-use form (8-7) and the modification-use form (8-8); i.e. from the pattern ca-ná:ki ‘mine’: am-ifa ‘my dog’ to the pattern ca-ná:ki ‘mine’: ifá ca-ná:ki ‘my dog’.84

The (b) pattern of spread of NP-use markers in Figure 7 shows the importance of N-based nominalizations, whose innovations spread to V-based nominalizations. Ryukyuan languages also illustrate this pattern of development. As noted above, the innovative replacement of the si marker by mun(u) in the NP-use of N-based nominalizations among Ryukyuan languages appears almost complete except for those peripheral islands mentioned above. This replacement pattern has spread to the NP-use of V-based nominalizations in a wide range of Ryukyuan languages except for the Okinawan dialects on the main island of Okinawa, where while mun(u) has replaced si in the NP-use of N-based nominalization, that of V-based nominalizations retain the older si-marker. Compare the following sets of data from Agena Okinawan and Irabu Ryukyuan of the Miyako archipelago in

84. A similar use of the noun meaning “thing” as an NP-use marker for N-based nominalizations is also seen in Thai, where the marker khɔːŋ with the lexical meaning of “thing” obligatorily marks NP-use. The spread of this marking to the modification context is again secondary in that it remains optional.
the south; in the latter the munu-marking has spread from the NP-use of N-based nominalization to that of V-based nominalizations.  

(8-10) Agena Okinawan (courtesy of Shigehisa Karimata)  
a. N-based nominalization  
i. ʔure: [sinsi: ga] mun=do:.
   that teacher NMLZR NPM=COP
   ‘That is the teacher’s.’  
ii. ʔure: [[sinsi: ga] kucu]=do:.
   that teacher NMLZR shoe=COP
   ‘Those are the teacher’s shoes.’  
b. V-based nominalization  
i. [ʔama=nkai tattʃo: (< -si ja)] wa:
   there-LOC stand.PROG.NPM.TOP (NM TOP) I.NMLZR
   ikigangwa.
   son
   ‘The one standing there is my son.’  
ii. [[ʔama=nkai tattʃo:-ru] ttʃo: (< ttʃu ja)] wa:
   there=LOC stand.PROG-NMLZR person.TOP person TOP I.NMLZR
   ikigangwa.
   son
   ‘The person who is standing there is my son.’  

(8-11) Irabu Ryukyuan (Miyako archipelago; Okinawa Prefecture, Japan; Shimoji 2008 and p.c.)  
a. N-based nominalization  
i. uri a [ba ga] munu.
   that TOP I NMLZR NPM
   ‘That is mine.’  
ii. uri a [[ba ga] zin]
   that TOP I NMLZR money
   ‘That is my money.’  
b. V-based nominalization  
i. [kama n taci-ur] munu u ba ga biki-vva.
   there LOC stand-PROG NPM TOP I NMLZR male-child
   ‘The one standing there is my son.’  

85. In addition to the munu-marked form in (8-10a), Irabu retains older nominalization forms involving the particle si as well as the even older verbal nominalizer, known as rentai-kei ‘adnominal form’ in Japanese linguistics.
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

ii. \[ ([kama n taci-ur] ffa) u ba ga
there LOC stand-PROG child TOP I NMLZR
biki-vva. (Modification-use)
\]
male-child
‘The child who is standing there is my son.’

Finally, we shall examine crosslinguistic marking patterns that show the development pattern (c) in Figure 7, namely the spread of a marker from the NP-use of V-based nominalization to the modification-use. Recall that Toba allows V-based nominalizations without any nominalization marker, which, however, must be marked by a demonstrative determiner when they head an NP, similar to the Portuguese pronoun-based nominalization in NP-use seen earlier. In the modification-use, however, a demonstrative determiner is not used, as in (8-12b) below.

(8-12) Toba (courtesy of Cristina Messineo)
a. NP-use
\[ [so [neta’age da Chaco]NMLZ]NP i-waGan so Juan
DD 3.exist.DIR DD Chaco 3-hit DD Juan
‘The one who lives in Chaco hit Juan.’
\]
b. Modification-use
\[ [[[so fijaGawa [Ø [neta’age da Chaco]NMLZ]NP i-waGan so Juan
DD man 3.exist.DIR DD Chaco 3-hit DD Juan
‘The man who lives in Chaco hit Juan.’
\]

The demonstrative determiner marking in NP-use has not yet been extended to the modification context above, where the Ø marker indicates its absence in (8-12b).86

When we turn to K’ichee’, we see that the determiner marking of V-based nominalization in NP-use has been extended to the modification context, as seen below.

(8-13) K’ichee’ (courtesy of Telma Can Pixabaj)
a. \[ x-Ø-inw-il lee [ixoq]
ASP-3SG.ABS-1SG.ERG-see the woman
‘I saw the woman.’
\]
b. \[ x-Ø-inw-il lee [x-Ø-u-ch’ay lee
ASP-3SG.ABS-1SG.ERG-see the/NMLZR ASP-3SG.ABS-3SG.ERG-hit the
achih]NMLZ
man
‘I saw the one whom the man hit/I saw the one who hit the man.’

86. Presumably it is possible to use a DD in the place of Ø in (8-12b), but it would result in a structure different from a restrictive relative clause construction. See Section 5.5.
c.  

\[
\begin{align*}
&x\text{-Ø-inw-}il \quad [\text{lee } \text{ixoq} \quad [\text{lee } \\
&\text{ASP-3SG.ABS-1SG.ERG-see} \quad \text{the woman} \quad \text{the/NMLZR} \\
&[x\text{-Ø-u-ch'ay} \quad \text{lee } \text{achih}]_{\text{NMLZ}}] \\
&\text{ASP-3SG.ABS-3SG.ERG-hit the man}
\end{align*}
\]

‘I saw the woman whom the man hit/I saw the woman who hit the man.’

The use of \textit{lee}, or its dialectal variants, in the modification context as in (8-13c) does not seem entirely obligatory at present, though its use appears highly favored according to the investigation of its status by a K’ichee’ specialist known to the present author. In other words, the determiner \textit{lee} is in a final stage of becoming a nominalizer/relativizer, so that V-based nominalizations become formally uniform in both the contexts of NP-use and modification-use, as in (8-13b) and (8-13c). Compare these with the Toba forms in (8-12), where the forms of V-based nominalizations are distinguished according to the usage pattern. K’ichee’ would eventually reach the stage where the determiner \textit{lee} becomes an obligatory nominalizer/relativizer as in German, which has developed nominalizers out of demonstrative pronouns. Observe the following where \textit{der} marking a V-based nominalization is obligatory in both NP-use (8-14a) and modification-use (8-14b).

(8-14) German

\begin{align*}
\text{a. } [\text{Der} & \quad [\text{der} \quad \text{morgen} \quad \text{kommt}]_{\text{NMLZ}}]_{\text{NP}} \quad \text{ist} \\
\text{art.msc.sub} & \quad \text{sub-nmlzr.msc tomorrow comes} \quad \text{is} \\
& \quad \text{mein Freund}. \\
& \quad \text{my friend} \\
& \quad \text{‘The one who (MSC) comes tomorrow is my friend.’}
\end{align*}

\begin{align*}
\text{b. } [\text{Der} & \quad \text{Mann} \quad [\text{der} \quad \text{morgen} \quad \text{kommt}]_{\text{NMLZ}}]_{\text{NP}} \quad \text{ist} \\
\text{art.msc.sub} & \quad \text{man} \quad \text{sub-nmlzr.msc tomorrow comes} \quad \text{is} \\
& \quad \text{mein Freund}. \\
& \quad \text{my friend} \\
& \quad \text{‘The man who comes tomorrow is my friend.’}
\end{align*}

The above discussions identify three sources of nominalization markers, namely (i) particles whose origins are unclear as in the cases of the Korean \textit{kes} and the nominalizers for N-based nominalization (aka genitive particles) \textit{no} (and \textit{ga}) in Japanese, (ii) a lexical source in terms of nouns meaning “thing” or others meaning “person” (Gã mɔ́ ’person’), “matter” (Japanese \textit{koto}, not discussed in this paper), “place” (Thai \textit{thít}), and (iii) determiners/demonstratives. At least some of these clearly start out as NP-use markers, and some of them have already become nominalizers, marking all occurrences of grammatical nominalizations as such.

Besides the nominalization markers above, many languages of the world use noun classifiers as makers of grammatical nominalizations, some of which have become nominalizers. First of all, the widely held understandings of classifiers...
miss the mark and fail to recognize their nominalization function.\textsuperscript{87} Even those who recognize the nominalizing function of classifiers do not appear to properly grasp its scope due to a limited understanding of what nominalization is all about.\textsuperscript{88} Because of this, we spend the next several pages detailing the nominalization function of classifiers and showing that classifier-marked nominalizations behave exactly like ordinary nominalizations in both NP- and modification-use. The marking patterns in these two functions also parallel between classifier-marked nominalizations and ordinary nominalizations, as we shall see. The essential difference between ordinary nominalizers and (nominalizing) classifiers is simply that the former derive nominals with a new denotation without classifying them, the latter both nominalize and classify derived nominals into different categories based on the nature of the new denotation, such as shape, size, consistency, animacy, gender, and function. The nominalization function of classifiers is clearly seen when they derive new nouns (lexical nominalizations), as in the Hmong and Barasano examples below.


a. V-based
   i. \textit{said} ‘see’ > \textit{tus} \textit{said} ‘supervisor’\textsuperscript{89}
   ii. \textit{sau} ‘write’ > \textit{tus} \textit{sau} ‘writer’

b. N-based
   i. \textit{dej} ‘water’ > \textit{tus} \textit{dej} ‘river’
   ii. \textit{ntawv} ‘paper’ > \textit{phau} \textit{ntawv} ‘book’
   iii. \textit{ntawv} ‘paper’ > \textit{tsab} \textit{ntawv} ‘letter’

\textsuperscript{87} Observe the following; “[a] classifier denotes some salient perceived or imputed characteristic of the entity to which an associated noun refers (or may refer).” (Allen 1977: 285). “Numeral classifiers are morphemes that only appear next to a numeral, or a quantifier; they may categorize the referent of a noun in terms of its animacy, shape, and other inherent properties.” (Aikhenvald 2006: 466) Numeral classifiers actually apply to verbs, nouns, demonstratives as well as grammatical nominalizations; see below). Compare these definitions with the following characterization of Barasano classifiers by Jones & Jones (1991: 49): “Barasano has an extensive system of noun classifiers, which provide concordance (agreement) within the noun phrase, and are used to form referring expressions that head noun phrases.” (Emphasis added).

\textsuperscript{88} Payne (1985: Chapter 4), who clearly recognizes the nominalization function of classifiers in Yagua, talks about the derivational (nominalizing) function and the inflectional function of classifiers, the latter of which is actually no more than modification-use of classifier-marked nominalizations (see (8-17b) and (8-18b)).

\textsuperscript{89} \textit{tus} = classifier for human beings, animals, things that closely affect people, and things that come in short length; \textit{phau} = stacks of things and things piled up on each other; \textit{tsab} = written messages.
(8-16) Barasano (Jones & Jones 1991)
   a. V-based
      i. bōa-gu `a male worker'
         work-prox.masc
      ii. bue-go `a female student'
         study-prox.fem
   b. N-based
      i. riri-rahe `clay cylinder (used as cooking pot support)'
         clay-cylindrical
      ii. kōbe-hāi `machete'
         metal-flat.thin
      iii. hēā-bu `firearm'
         fire-cylindrical container
   'firearm'

So-called numeral classifiers are numeral-based grammatical nominalizations that derive nominal structures denoting entities characterized by quantity and then classified. E.g.,

(8-17) Thai (Iwasaki & Ingkaphirom 2005: 69, 75)
   a. NP-use
      [sām tua] nān hāay pay nāy
      three clf that disappear.asp where
      'Where have the three (e.g. dogs) gone?'
   b. Modification-use
      nī khāay [māa [sām tua]]
      Ni sell dog three clf
      'Ni sold three dogs.'

(8-18) Ainu (Bugaeva 2012: 811)
   a. NP-use
      cōka [tu-n] ci=ne na, [tu-p] en=kor-e
      1pl.exc two-hum 1pl.(exc).a=cop fin two-thing 1sg.o=have-caus
      yan
      imp.pol
      'There are two of us (lit. 'we are two humans'), so give us two pieces.'
b. Modification-use\textsuperscript{90}

i. \([menoko \ {tu-n}]\)
   \begin{itemize}
   \item woman \ two-HUM
   \end{itemize}
   ‘two women’

ii. \([chise \ {tu-p}]\)
   \begin{itemize}
   \item house \ two-THING
   \end{itemize}
   ‘two houses’

(8-19) Barasano (Jones & Jones 1991: 50, 59, 113)

a. NP-use

\begin{itemize}
\item \(yu\text{ku} \ ha \ su\text{re} \ use \ yu\ {idia-ro}\)
\item Sabîdo \ bararî \ Luka
\item tree \ hit \ cut \ off \ many \ weave \ 1s \ three-clf \ Sabino \ four \ (fences) \ Luke
\item [\(h\text{ua-se}\)] \ Arike \ [\(h\text{ua-se}\)] \ use-ka-hu \ y\text{u}\a
\item two-clf \ Eric \ \two-clf \ weave-far.pst---3 \ 1x
\end{itemize}

‘Cutting down trees, weaving, I (wove) three (length of) fences, Sabino (wove) four, Luke (wove) two, (and) Eric wove two.’

b. Modification-use

i. \([h\text{uu-re} \ [h\text{u-u-re}]]\)
   \begin{itemize}
   \item \(\text{h}o\text{bo-a-ha}\)
   \item two-clf-obj \ want-pres---3 \ I
   \end{itemize}
   ‘I want two hammocks.’

ii. \([h\text{ua-rahe} \ [h\text{ua-rahe}]]\)
   \begin{itemize}
   \item \(k\text{obe-prahe-ri}\)
   \item two-cylindric \ metal-cylinder-pl
   \end{itemize}
   ‘two metal cans’

As already seen, classifiers also nominalize pronouns, ordinary nouns and noun phrases giving rise to so-called “relational/possessive” classifiers (Aikhenvald 2006), which are actually nominal-based nominalizations known as the genitive/possessive form discussed in Section 7. The following Cantonese example shows both modification- and NP-use of N-based nominalizations marked by a classifier.

(8-20) Cantonese (courtesy of Haowen Jiang)

\begin{itemize}
\item \(keoi^5 \ gaan^1 \ fong^2\)
\item \(daa^6 \ gwo^3\)
\item \(ngo^{23} \ gaan^1\)
\item 3SG \ CLF \ room \ big \ exceed \ I \ CLF
\end{itemize}

‘His room is bigger than mine.’

As below, possessions are marked by classifiers in Barasano, and these can also have a modification-use (8-21b), though the head nouns are typically unexpressed since the classifiers and context sufficiently narrow down the denotation/reference of the entire noun phrase (Jones and Jones 1991: 62).

\textsuperscript{90} In prenominal position, numerals may directly modify the head noun; e.g. \(sine \ menoko\) ‘one woman’. 
(8-21) Barasano (Jones & Jones (1991: 62))

a. i. yū ya-ro
   1SG NMLZR-CLF
   ‘mine’ (e.g. seat)

ii. ĩ ya-ga
    3MSG NMLZR-HOLLOW
    ‘his hollow thing,’ ‘his thing having a hole’ (e.g. needle)

iii. so ya-tuti
     3FSG NMLZR-STACK
     ‘her stack’ (e.g. book)

b. i. [ĩ ya-gw] jū
    he NMLZR-MASC hammock
    ‘his hammock’

ii. [yū y-u] bāk-u
    I NMLZR-MASC progeny-MASC
    ‘my son’

Classifiers may also derive new nominal structures from grammatical nominalizations. The Thai argument nominalization (8-22a) below may denote/refer to all kinds of things, including a cat, a shirt, a skirt, a desk, a knife, a notebook, as well as a book.91

(8-22) Thai (courtesy of Kingkarn Thapkanjana)

a. [thīi chán suāu mūawaannii] pheeŋ māk
   NMLZR I buy yesterday very expensive
   ‘What I bought yesterday was very expensive.’

b. [tua [thīi chán suāu mūawaannii]] pheeŋ māk
   CLF NMLZR I buy yesterday very expensive
   ‘What I bought yesterday was very expensive.’

c. [lēm [thīi chán suāu mūawaannii]] pheeŋ māk
   CLF NMLZR I buy yesterday very expensive
   ‘What I bought yesterday was very expensive.’

The form in (8-22b) with the classifier tua (< ‘body’) is a new nominal structure that denotes only things such as a cat, a shirt, and a skirt, a desk, which have some sense of association with a body. On the other hand, (8-23c) with lēm refers to sharp, slender objects and book-like bound objects, such as a knife, a book and a notebook. This pattern parallels that found in the classifier marking of N-based

---

91. Cf. Hmong pattern: [uas dawb] yog kuv tus ([NMLZR white] is I CLF) ‘The one (e.g. skirt, pencil, etc.) that is white is mine’, daim [(uas) dawb] yog kuv tus (CLF [NMLZR white] in I CLF) ‘The one (e.g. skirt, ‘pencil) that is white is mine’, tus [(uas) dawb] yog kuv tus ‘The one (e.g. pencil, ‘skirt) that is white is mine’ (curtesy of Martha Ratliff).
nominalizations; \([khɔŋɔ chán]\) 'mine (all kinds of things)', \(tua [khɔŋɔ chán]\) 'mine (e.g. a cat, a shirt, a skirt, a desk)', \(lɛm [khɔŋɔ chán]\) 'mine (e.g. a knife, a book, a notebook)'.

A similar pattern is also seen Barasano, as below, where classifier marking derives a structure denoting more specific objects out of a grammatical nominalization with a general denotation.

(8-23) Barasano (Jones & Jones 1991: 146)

a. \(ĩ\ 3ms\ su-a-ka-ti\)
   \[3ms\ weav-e-far.pst-\--prox.nmlzr\]
   'things he wove'

b. \(ĩ\ 3ms\ su-a-ka-ti-bu\)
   \[3ms\ weav-e-far.pst-\--prox.nmlzr-\--basket\]
   'the basket (I saw) him weave' or 'his woven basket'

Classifiers, thus, play various nominalizing functions across languages. On the other hand, the use of classifiers differs considerably from one language to another but a general trend is seen, similar to the development of nominalization markers examined above. In general, the pattern of classifier marking in the modification context is equivalent to or less developed/grammaticalized than that in the context of NP-use. In Thai classifiers are generally "optional" in the sense that a given form may or may not be marked by a classifier, though the use of a classifier changes meaning (cf. (8-22)). A major exception is numerals that need to be nominalized by a classifier in order to function as an entity-denoting nominal (as opposed to denoting numbers and numerals). Compare the numeral-based nominalizations, where classifier marking is obligatory in both NP- and modification-use, and the demonstrative-based ones below, in which classifier marking is optional in both contexts.

(8-24) Thai (cf. (8-17))

a.  
   \[sâm \*(tua)] nán hâay pay nây\]
   three CLF that disappear ASP where
   'Where have those three (e.g. dogs) gone?'

b. Modification-use
   \[ni khâay [màa [sâm \*(tua)]]]
   Ni sell dog three CLF
   'Ni sold three dogs.'

92. The suffix \(-ti\) is a time-bound nominalizer associated with the far past tense marker \(-ka\).

93. The notation \(\*\*(...)\) indicates an obligatory element, as opposed to \(\(\ldots\)\) indicating an optional element.
Thai (Iwasaki & Ingkaphirom (2005: 64) and Shoichi Iwasaki p.c.)

a. NP-use

\[
\begin{align*}
\text{[(khan) nii] thawrây} \\
\text{clf this how.much} \\
\text{‘How much is this?’, ‘How much is this clf (e.g. car)?’}
\end{align*}
\]

b. Modification-use

\[
\begin{align*}
\text{rót [(khan) nii] thawrây} \\
\text{car clf this how.much} \\
\text{‘How much is this car?’}
\end{align*}
\]

Unlike Hmong and Cantonese discussed below, Thai does not use a classifier to nominalize pronouns and ordinary nouns in forming so-called possessive constructions. Instead, they have recruited the noun \textit{kʰɔ̌ɔŋ ‘thing’}, which, when used as the nominalizer for N-based nominalizations, is grammaticalized to the extent that it can denote people and animals as well. This nominalizer is obligatory in the NP-use of N-based nominalizations, but optional in the modification-use, as shown below.

Thai (Iwasaki & Ingkaphirom (2005: 13, 69)

a. NP-use

\[
\begin{align*}
\text{[*(kʰɔ̌ɔŋ) khun] yău năy} \\
\text{nmlzr you stay where} \\
\text{‘Where is yours?’}
\end{align*}
\]

b. Modification-use

\[
\begin{align*}
\text{náŋs [((kʰɔ̌ɔŋ) nók]} \\
\text{book nmlzr Nok} \\
\text{‘Nok’s book.’}
\end{align*}
\]

In the case of Thai V-based argument nominalizations, a classifier is optional in both NP- and modification-use. Thus, parallel to the NP-use pattern seen in (8-22) above, a classifier may or may not occur in the modification context, as below.

Thai

a. \textit{krâprooŋ [(tua) [thîi chán suáu mûawaannīi]] pheëŋ mâk}

\[
\begin{align*}
\text{skirt clf nmlzr I buy yesterday very expensive} \\
\text{‘The skirt that I bought yesterday was very expensive.’}
\end{align*}
\]

b. \textit{nâŋsùu [(lêm) [thîi chán suáu mûawaannīi]] pheëŋ mâk}

\[
\begin{align*}
\text{book clf nmlzr I buy yesterday very expensive} \\
\text{‘The book that I bought yesterday was very expensive.’}
\end{align*}
\]

White Hmong uses classifiers obligatorily in marking not only numerals, but also demonstratives and N-based nominalizations (possessive constructions).
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

(8-28) White Hmong (courtesy of Martha Ratliff)
   a. *Kuv muaj [peb tus]
      I have three CLF
      ‘I have three of them (e.g. pencil).’
   b. *Kuv muaj [tus no]
      I have CLF this
      ‘I have this one (e.g. pencil).’
   c. *Kuv muaj [kuv tus]
      I have I CLF
      ‘I have mine (e.g. pencil).’

Like all these forms, V-based grammatical nominalizations in Hmong need to be marked by a classifier in NP-use even when marked by the nominalizer uas. For example;

(8-29) White Hmong (courtesy of Nerida Jerkey)
   [*(tus) [uas hais]] yog Maiv Yaj
      CLF NMLZR say COP Mai Ya
      ‘The one who told (this story) is Mai Ya.’

The above fact is consistent with the fact that in Hmong classifiers function like a determiner marking all definite or specific nouns. Grammatical nominalizations in modification-use, however, cannot be marked by a classifier, and the following is ungrammatical; removing the second classifier daim renders the sentence grammatical.

(8-30) White Hmong (courtesy of Martha Ratliff)
   *Daim tiab [daim [uas kuv niam ntxhua]] yog kuv daim.
      CLF skirt CLF NMLZR I mother wash COP I CLF
      ‘The skirt that my mother is washing is mine.’

Notice that the nominalizer uas is ommissible in White Hmong under certain circumstances, sometimes preferably, e.g. when short nominalizations are involved.

(8-31) White Hmong (courtesy of Martha Ratliff)
   a. [Daim [uas dawb]] yog kuv daim.
      CLF NMLZ white COP I CLF
      ‘The white one (e.g. a skirt) is mine.’
   b. [Daim [dawb]] yog kuv daim.
      CLF white COP I CLF
      ‘The white one is mine.’
      CLF skirt CLF white COP I CLF
      ‘The white skirt is mine.’
While in (8-31b) the classifier daim appears to be functioning like a nominalizer, a verbal nominalization marked by it cannot appear in a modification context, as shown in (8-31c), which must replace the classifier daim with the nominalizer uas to become grammatical.

The non-verbal based nominalizations marked by classifiers (see (8-28) above) may modify a nominal. While some report (e.g. Nerida Jerkey, p.c.) tells us that they cannot modify those that are already marked by a classifier, as in (8-32a)–(8-32c), our own Hmong consultant permits doubling of classifier marking, as in (8-32d)–(8-32e).

(8-32) White Hmong
a. [[ob lub] [(lub) rooj]]
   two CLF CLF table
   ‘two tables’

b. [[kuv lub] [(lub) rooj]]
   I CLF CLF table
   ‘my table’

c. [[kuv (lub)] [ob lub rooj]]
   I CLF two CLF table
   ‘my two tables’

d. [[daim nplooj] [(daim) no]]
   CLF leaf CLF this
   ‘this leaf’

e. [[kuv daim] [(daim) ntawd]]
   I CLF CLF that
   (lit.) ‘my that one’, ‘that one of mine’

In any event, it is clear that classifier-marked nominalizations are more limited in the modification context than in the NP-use context, where classifier marking is obligatory.

Compared to Thai and Hmong situations above, Cantonese classifiers may mark grammatical nominalizations very consistently in both NP- and modification-use, showing that in this language classifiers are beginning to supplant the general nominalizer ke33 (cf. (b) and (c) below), corresponding to the Mandarin nominalizer de.

(8-33) Cantonese (courtesy of Haowen Jiang; Matthews & Yip 1994: 111, 112)
a. ngo23 kin33 tou35 sam55 pun35 (fy55)
   I see reach three CLF book
   ‘I see three (books).’
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

143

b. $\text{ni}^{55} \text{pun}^{35} (\text{fy}^{55}) \text{hei}^{22} \text{ngo}^{23} \text{ke}^{33}$
   this CLF book COP I NMLZR
   ‘This (book) is mine.’

c. $\text{keoi}^{5} \text{gaan}^{1} (\text{fong}^{2}) \text{daai}^{6} \text{gwo}^{3} \text{ngo}^{23} \text{gaan}^{1}$
   3SG CLF room big exceed I CLF
   ‘His (room) is bigger than mine.’

d. $\text{[[gaau}^{3} \text{lei}^{5} \text{taan}^{4} \text{kam}^{4}] \text{go}^{5}? \text{teach} \text{you} \text{play} \text{piano} \text{that} \text{CLF}$
   ‘The one who teaches you piano?’

d’. $\text{[[ngo}^{5} \text{dei}^{5} \text{hai}^{2} \text{faat}^{4} \text{gwok}^{3} \text{sik}^{6}] \text{di je}^{5} \text{gei}^{2} \text{hou}^{2}\text{-sik}^{5}]$
   we in France eat CLF food quite good-eat
   ‘The food we ate in France was pretty good.’

Classifiers have not yet taken over the general nominalizer $\text{ke}^{33}$ in Cantonese, where the latter is still a default nominalizer. The use of classifiers in nominalization function, except for numerals and demonstratives, is said to be part of colloquial speech and there are also contexts in which classifiers and $\text{ke}^{33}$ are not interchangeable. For example, $\text{ke}^{33}$ in (8-33b) cannot be replaced by $\text{pun}^{35}$; but see (8-33c), where the classifier marking $\text{ng}^{2} \text{gaan}^{1}$ (I CLF) ‘mine’ is permitted. Like Hmong above, Cantonese also disfavors the use of a CLF-marked modifier when the head nominal is also marked by a classifier. Thus, while the classifier-marked modifier is permitted in (8-33b), (8-33c) above, (8-34a) below is not possible, requiring a modifier marked by the default nominalizer, as in (8-34b).94

(8-34) Cantonese (courtesy of Haowen Jiang)

a. *$[[\text{ng}^{23} \text{pun}^{35}] [[\text{ni}^{55} \text{pun}^{35}] \text{fy}^{55}]]$95
   I CLF this CLF book
   (lit.) ‘my this book’, ‘this book of mine’

b. $[[\text{ng}^{23} \text{ke}^{33}] [[\text{ni}^{55} \text{pun}^{35}] \text{fy}^{55}]]$
   I NMLZR this CLF book
   (lit.) ‘my this book’, ‘this book of mine’

94. According to Haowen Jiang (p.c.), there are other southern Chinese dialects (e.g. Jixi Huitseu Mandarin) in which classifier marking is more advanced than in Cantonese, allowing forms that are not possible in the latter. The crosslinguistic pattern of classifier-marking seems to suggest the following hierarchy, where the numerals are the the easiest to nominalize by classifiers: NUM > DEM > V-based/(Pro)NOUN NMLZ.

95. The avoidance of the use of classifiers in this modification context is reminiscent of the prohibition of the use of a plural form of V-based nominalizations in the modification context when the head noun is marked plural (see (6–27)).
In contrast to Hmong and Cantonese, the Newar animate (-mha) and inanimate (-gu:) classifiers have fully given rise to the nominalizers -mha and -gu (NB, a short u), and the classifiers (e.g. -gu:) now nominalize only numerals. That is, in Modern Newar only numerals are nominalized by -gu: and other appropriate classifiers (e.g. -ga: for cars), and all else must use the nominalizer -gu, as below:

(8-35) Newar (courtesy of Kazuyuki Kiryu)

a. \( ji: \) cha-gu/ni-gu: (saphu:) khane du
   1sg.erg one-clf/two-clf book see can.nd
   ‘I can see one/two (books).’

b. \( twā\)\(^{96}\) (saphu:) bwā
   this book read.imp
   ‘Read this (book).’

b’. \( thu=\)gu (saphu:) bwā.
   this=nmlzr book read.imp
   ‘Read this (book).’

c. wa gāri Rām=yā=gu (kha:)
   that car Ram=nmlzr=nmlzr cop
   ‘That car is Ram’s.’

c’. wa Rām=yā=gu (kha:)
   gāri (kha:)
   that Ram=nmlzr=nmlzr car cop
   ‘That is Ram’s car.’

d. \( thu: \) [tuyu=gu (wāsa:)] phi: nu
   today white=nmlzr clothes wear.fc hortative
   ‘Today, let’s wear the white (clothes).’

As can be seen in (8-35c’), nominalizer marking is optional in the modification-use of N-based nominalizations in Newar, which is marked by the N-based nominalizer yā.

Classifier languages of South America generally show classifier usage-patterns similar to the Asian counterarts studied above. In particular, many languages use classifiers as nominalizers with some restrictions of their use in the modification context. Yagua, a Peba-Yaguan language spoken primarily in northeastern Peru, obligatorily mark numerals and demonstratives with classifiers in both NP- and modification-contexts, while the marking of adjective-based nominalizations is optional in the modification context. Observe:

(8-36) Yagua (Payne 1985: 131, 174, 176, 179, 189)

a. … tāáyu tā-duu-guui-dee-tēe vurya-jiriy-tēe
   why one-tube-one-dim-emph we.incl-grab-emph
   ‘…why did we get one (flute)?’

\(^{96}\) This is likely to be an older nominalized form.
As shown in (8-36a, a’), Yagua infixes a classifier for numeral-based nominalizations in both NP and modification uses. Example (8-36b″) shows that classifier marking is retained in the modification context for demonstrative-based nominalizations. Notice further that modification by a classifier-marked demonstrative can modify a classifier-marked head, which was optional in Hmong and was avoided in Cantonese, showing that leveling of the demonstrative forms in the two use contexts is more advanced in Yagua. The usage pattern of adjective-based nominalizations by classifiers differs slightly from this. That the NP-use of adjective-based nominalizations requires classifier marking, like numeral-based and demonstrative-based forms, is clearly indicated by the ungrammatical form in (8-36c) and the grammatical form in (8-36c’). In the modification context, adjective-based nominalizations may be marked by a classifier, as in (8-36c″), or may not, as in (8-36c‴), indicating that classifier marking for adjective-based nominalizations in Yagua is not fully grammaticalized in the modification context.
As for V-based nominalizations, Yagua involves a nominalization marker, which typically consists of a demonstrative, a classifier, and the ending -tìy. Observe:

(8-37) Yagua (Payne 1985: 106, 109)
a. $naan$-$q$ $juno$-$râ$ $[jiy$-$ra$-$tìy$ $rà$-$ranîy]$  
$\text{1DLLEXICL-IRR CUTF-INAN DEM-CLF.NEUT-TIY INAN-stand}$  
‘We are going to cut this which is standing.’
b. $ray$-$mutiyey$-$râ$ $jimyichara$ $[jiy$-$ra$-$tìy$ $sa$-$tàâryuy]$  
$\text{1SG-COOK-INAN FOOD DEM-CLF.NEUT-TIY 3SG-BUY}$ $Tomâsa-râ$  
$Tom-INAN$  
‘I cooked the food that Tom bought.’

Besides classifier-marked nominalizers involving demonstratives, Yagua seems to allow the nominalizer ending -tìy to attach directly to a pronoun (see Payne (1985: Chapter 4) for such examples).

Barasano, an Eastern Tocanoan language spoken in southeastern Colombia, uses classifying nominalizers very extensively. Unlike Yagua, where N-based nominalizations (so-called possessive/genitive constructions) are not marked by classifiers, and where V-based nominalizations (so-called relative clauses) involve a special marker, Barasano uses classifiers, which are numerous in number and which vary in form, to mark all these types of nominalization. Classifier-marked numerals denote objects classified according the classifier marking them, as in (8-38a). These can be used as modifiers as in (8-38a’). But many classifiers being specific enough, modification structures like these are far less commonly used; the NP-use of classifier-marked numerals does the job. That is, instead of (8-38a’.ii), the one without the head noun is most likely used (cf. (8-38c.iii)).

(8-38) Barasano (Jones & Jones 1991: 50, 57, 60, 61, 62, 66, 78, 86)
a. i. $hua$-$hâi$  
$\text{TWO-FLAT.THIN}$  
‘two flat, thin objects’
ii. $hua$-$râ$  
$\text{TWO-ANM.PL}$  
‘two living things’
a’. i. $hua$-$hâi$ $sudi$-$hâi$  
$\text{TWO-FLAT.THIN CLOTH-FLAT.THIN}$  
‘two pieces of cloth’
ii. $hũn$-$re$ $hũ$-$n$-$re$ $ābo$-$a$-$ha$ $yu$ $hammock$-$OBJ$ $TWO$-$HAMMOCK$ $WANT$-$PRS$--$3$ $I$  
‘I want two hammocks.’
b. i. adi-re ābi-a-bu yu
these-obj pick.up-mot-pst~3 I
‘I brought these.’

ii. adi-tuti-re ābi-a-ha bu
this-stack-obj pick.up-mot-fut.impv.prox you
‘You take this stack.thing (book).’

b’. i. ti gūbu-re buha …
that log-obj find
‘finding that log’

ii. nye-godo-a ti-a-re ābo-a-ha yu
oil-cleared.out-hollow that-hollow-obj want-prs~3 I
‘I want that empty oil can.’

c. i. yu ya-ro
I NMLZR-THING
‘my thing (e.g. seat)’

ii. ī ya-ga
he NMLZR-HOLLOW
‘his hollow thing/thing having a hole (e.g. needle)’

iii. haihe ī ya-tuti yā-a-ha ti
Jim he NMLZR-STACK be-prs~3 it
(lit.) ‘That is Jim’s (stack stuff).’ ‘That is Jim’s book.’

c’. hū ī ya-gu
hammock he NMLZR-HAMMOCK
‘his long hammock’

As in (8-38b, b’) above, a demonstrative may or may not be marked by classifiers in both NP- and modification-use, whereas N-based nominalizations require classifier marking in both NP- and modification-use, as in (8-38c, c’).

V-based nominalizations in Barasano are formed in two ways. A first method is to attach nominalizing classifiers directly to verb roots and stems as in the examples below, involving the general classifiers -se and -ro, meaning “thing”, “time”, “place”, and time-bound nominalizer (-ka-ti ‘far past, non-proximal’), as well as more specific nominalizing classifiers that distinguish animacy-number (-rã ‘inanimate singular’) and gender-number (-go ‘feminine singular’, -gu ‘masculine singular’).97

(8-39) Barasano (Jones & Jones 1991)

a. wu-se
fly-clf
‘flying things (airplanes)’

97. Some of these can be further marked by a classifier; e.g. compare (8-40b) and yā-ro-hū (be-nmlzr-clf.place) ‘the place where’.
b. \( i\ y\-\text{ro} \)  
he be-clf  
‘when/where he is’
c. \( i\ s\-\text{uka-ka-ti} \)  
he weave-far.pst-prox.nmlzr  
‘things he wove’
d. \( b\-\text{ue-ria-r\-\}} \)  
study-pst.conj-anp  
‘ones who probably studied’
e. i. \( b\-\text{oa-g\-u} \)  
work-masc.sg  
‘a male worker’
ii. \( b\-\text{ue-go} \)  
study-fem.sg  
‘a female student’
f. \( i\ s\-\text{uka-ka-ti} \)  
he weave-far.pst-prox.nmlzr  
‘things he wove’

All these allow both NP- and modification-use, as illustrated below.

(8-40) NP-use of V-based nominalizations (Jones & Jones 1991: 85, 148, 150, 169)
a. \( k\-\text{ahi} [i-d\-\text{i-re}] i-d\-\text{i-ka-ti} b\-\text{u} \)  
coca drink-clf drink-far.pst-q you  
‘Did you drink the halcinogenic drink?’
b. \( i\-\text{re} [\text{yu} i\-\text{si-boa-ka-ti-re}] \)  
he-obj I give-but-far.pst-prox.nmlzr-obj  
move.away-ff-infer-3 msg  
‘He threw away what I had given him.’
c. \( [\text{yu} y\-\text{i-boa-r\-\}, y\-\text{u-re} k\-\text{udi-beti-s-\-\text{a-\-id\-\}}}] \)  
say-but-anp I-obj respond-neg-prs.prox-3p they  
‘The ones I talked to aren’t responding to me.’
d. \( [a-d\-\text{o} e\-\text{ha-go-re}] b\-\text{asi-be-a-ha} y\-\text{u} \)  
here arrive-fem.sg-obj know-neg-prs.prox--3 I  
‘I don’t know the feminine one who just arrived here.’

(8-41) Modification-use of V-based nominalizations (Jones & Jones 1991: 6, 143, 151)
a. \( [o\-\text{ko} [k\-\text{edi-se}]] \)  
water fall-clf  
(lit.) ‘falling-thing water’, ‘rain’
b. \( t\-\text{o k\-\text{o-ro} y\-\text{a-ka-b\-a} [r\-\text{obi-a [ue-o-ro]]} \)  
that count-nmlzr be-far.pst-3p female-p dirty-caus-anp  
‘There were that many women who were dirting (the river).’
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

149

c. \([gubo\ sudi\ tēdi\ [bū-re\ yu\ sēdi-ka-ti]]\)
foot\ clothes\ tennis.shoes\ you-obj\ I\ ask-FAR.PST--PROX.NMLZR
huar-ri\ buh
pick.up-q\ you
‘Did you get the tennis shoes that I asked you for?’

d. \([sīg-o\ [īdā\ rāka\ bue-go]]\ yu-re\ ăbo-a-bō\ so\ one.AN-3FEM.SG\ they\ with\ study-FEM.SG\ I-obj\ wantPRS-FEM.SG\ she
‘A girl who studies with them wants me (for her husband).’
(lit.) ‘A female who is a female student with them, she wants me.’

A second way to form V-based nominalizations in Barasano involves the suffix -ri, which is recognized as a nominalizer but is glossed as PTCPLE (participle) by Jones & Jones (1991: 43). While the other nominalizers discussed above appear to derive nominalizations with concrete denotations like things and persons, -ri derives event nominalizations as well, which is probably why Jones and Jones (1991) glosses it as PTCPLE. The event-denoting -ri nominalizations are widely used as adverbials including the use as chain structures discussed in Section 5.1. When they derive forms denoting concrete objects, their NP-use typically calls for classifier marking, though it may be possible to use them without a classifier, as in the following example.

(8-42) Barasano (Jones & Jones 1991: 43)
\([uka-ri\] kuti-go yā-ka-bō\)
write-NMLZR\ have-FEM.SG\ be-FAR.PST-3FEM.SG
‘She had spots (as a characteristic).’ (lit.) ‘She was a female possessor of writings/a female one who had writings.’

The more commonly observed usage pattern of -ri nominalizations has additional classifier marking as in the following examples.

(8-43) Barasano (Jones & Jones 1991: 24, 135, 150)
a. \([sūbe-ri-hāi]\ [sūa-ri-hāi]\ ăbo-a-ha\)
green-NMLZR-FLAT.THIN\ red-NMLZR-FLAT:THIN\ want-PRS--3
‘I want either a green cloth or a red one.’
b. \([yu\ īa\ buha-ri-hāi]\ yi-beti-busa-a-ha\ ti\)
I\ see\ find-NMLZR-FLAT:THIN\ dark-NEG-VERY.much-PRS--3\ it
‘The cloth I found isn’t very dark.’ (lit.) ‘What flat, thin thing I found, it isn’t very dark.’
c. \([ō\ kāhi-ri-ku]\ ăbo-a-ha\ yu\)
there\ hang-NMLZR-HAMMOCK\ want-PRS--3\ I
‘I want that hammock which is hanging there.’ (lit.) ‘I want that hanging hammock-thing.’
As expected, all these forms allow a modification-use as well, as below.

(8-44) Barasano (Jones & Jones 1991: 63, 66, 150)
   a. \[hũ \quad [ō \quad kāhi-ri-ku] \quad ābo-a-ha \quad yu\]
      hammock there hang-NMLZR-HAMMOCK want-PRS-~3 I (cf. (8-43c))
      ‘I want that hammock which is hanging there.’
   b. \[gahe \quad [gũbu \quad [yao-ri-ku]]\]
      another log long-NMLZR-LOG
      ‘another long log’
   c. \[sũa-ri-hãi \quad [yoa-ri-hãi]] \quad ābo-a-ha \quad yu\]
      red-NMLZR-FLAT.THIN long-CLE.FLAT.THIN want-PRS-~3 I
      ‘I want a long red piece of cloth.’ (lit.) ‘I want a red thin, flat thing that is long thin, flat.’

Interestingly and most relevant to the main point of the discussion in this subsection, -ri nominalizations, which appear to be normally marked by a classifier in NP-use, need not be marked by a classifier in the modification context, as below, showing that the marking pattern in the modification context lags behind that in the NP-use context.

(8-45) Barasano (Jones & Jones 1991: 21, 111, 144, 152)
   a. \[gahe \quad gũbu \quad [[yoa-ri] \quad gũbu]\]
      other log long-NMLZR log
      ‘another long log’
   b. \[[bũa-ri] \quad bãs-o\]^98
      work-NMLZR human-FEM.SG
      (lit.) ‘working female human,’ ‘a woman/girl worker’
   c. \[[sũa-ri] \quad bãs-ũ] \quad yã-a-bĩ \quad i\]
      kill-NMLZR human-MASC.SG be-PRS-3MASC.SG he
      (lit.) ‘He is a killing/killer male human.’ ‘He is a killer.’
   d. \[[iũ-ri] \quad hai-gu] \quad yã-a-bĩ\]
      dirty-NMLZR big-MASC.SG be-PRS-3MASC.SG
      (lit.) ‘He is a dirty big male.’ ‘He is very dirty.’ ‘He has a lot of dirt.’
      Cf. \[[hai-gu] \quad yã-a-bĩ\]
      big-MASC.SG be-PRS-3MASC.SG
      (lit.) ‘He is a big male.’ ‘He is big.’

^98. There is some issue regarding identification of the head in these forms, but in view of the author’s observations that “[t]he descriptive modifier generally precedes the head noun in the noun phrase” (Jones & Jones 1991: 4), the literal readings we provide seem to be more faithful interpretations of the data than the idiomatic translations by the authors. Barasano is an OV and postpositional language.
Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

Tuyuca, closely related to Barasano, has a classifier system similar to Barasano. According to Jane Barnes (p.c.), classifier marking is obligatory when N-based nominalization (aka gentives/possessives) is used as an NP-head, but it is optional in the modification context, similar to the pattern seen in Newar (see 8-35c, c'). For example,

(8-46) Tuyuca (Jane Barnes, p.c.)

a. \([yii \ paki-ya-ro]\)
   my father-NMLZR-2D.FLEXIBLE
   ‘my father’s’ (as in “They are my father’s/ My father’s are those.”)

b. \([\{yii \ paki-ya(-ro)\} \ sirir\}\]
   my father-NMLZR-2D.FLEXIBLE trouser
   ‘my father’s trousers’

The above study of classifier marking on nominalization structures in the two usage contexts of NP-use and modification-use also corroborate the pattern of development of nominalization markers recognized by Shibatani and Shigeno (2013) (see Figure 7 (p. 129). Specifically, in both N-based and V-based nominalizations, nominalization markers start out in NP-use, as markers of referential use of nominalizations. As in (a) and (c) in Figure 7, these markers then spread to the modification context, eventually becoming nominalizers, marking nominalization structures as such regardless of their usage contexts.99 The observed patterns of development of nominalization markers allow us to draw the following generalizations.

(8-47) Generalizations on the forms of nominalizations

The form of nominalizations in modification-use is equal to or simpler in formal complexity than that in NP-use. In particular, if a modifying nominalization involves nominalization morphology, the same marker is found in its NP-use, either in the contemporary language or in a historically earlier form or in its dialects.

The caveat about a historically earlier form of the language or its dialects in the above generalization is necessary since a new marker is likely to be introduced first in the context of NP-use either replacing the old marker or incrementally, once the uniformity in form is achieved in the two usage contexts.

We might be able to draw some kind of generalization from the pattern of development depicted as (b) in Figure 7, but since what is involved here are two different types of nominalization, N-based and V-based, rather than two different uses of a single type of generalization, it might be difficult to do so. In many

99. A complete study of the forms of nominalizations must examine the marking patterns in nominalizations in the adverbial function as well.
languages N-based and V-based nominalizations have distinct morphology, and we need more cases attesting the pattern of development depicted in Figure 7(b).

9. Summary and implications

After summarizing the discussions above in general terms, the final Subsection 9.1 discusses the implications of the present study for both descriptive and theoretical studies.

Past studies on nominalization tended to focus on lexical nominalizations because they typically have clear morphological marking. We showed in the beginning that across different languages the same lexical nominalization morphology may apply to units larger than words, suggesting the existence of grammatical nominalizations. The field has been slow to recognize grammatical nominalizations because many do not have clear nominal morphology or the forms involved have the same verbal form as clauses and sentences. We have argued that the notion of nominalization is neither morphological nor syntactic, but functional. Crosslinguistic investigations reveal clearly that formally different structures cohere in their semantics and usage patterns, supporting this view. At the same time, such studies provide crucial evidence, including morphological support, that is hard to find when dealing with single languages such as Japanese and English.

We have argued that traditional studies fail to distinguish between structures and their use, a failure that has led to the recognition of what is no more than different uses of the same basic structures as independent constructions. We have argued strongly that relative clauses are simply uses of grammatical argument nominalizations. So-called internally-headed RCs are event nominalizations in NP-use, which evoke various concepts metonymically related to the events such as the abstract concepts of facts and propositions or concrete concepts such as event protagonists and resultant products. What are known as headless relative clauses are instances of the NP-use of argument nominalizations, which also have a modification-use giving rise to what are known as relative clause constructions with a modified head noun.

The reanalysis of the genitive case or the possessive form as a nominal-based nominalization reveals that nothing like possessive pronouns and possessive adjectives exist as separate parts of speech. Similar to the case of relative clauses, they are no more than two uses of N-based nominalizations. We have provided ample crosslinguistic evidence in support of this new analysis, including classifier marking, which nominalizes and then classifies the entities denoted.

Table 2 below shows how disparate constructions and those thought to be related yet independent construction types in the past studies receive a uniform
analysis in terms of the notion of grammatical nominalizations of different types under different usage patterns:

Table 2. Comparison of past treatments and nominalization-based analysis

<table>
<thead>
<tr>
<th>Past treatments</th>
<th>Nominalization-based analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Verb complement</td>
<td>Event nominalization: NP-use</td>
</tr>
<tr>
<td>b. Noun complement</td>
<td>Event nominalization: (Adnominal) Modification-use</td>
</tr>
<tr>
<td>c. Internally-headed RC</td>
<td>Event nominalization: NP-use</td>
</tr>
<tr>
<td>d. Headless/free RC</td>
<td>Argument nominalization: NP-use</td>
</tr>
<tr>
<td>f. Relative clause (RC)</td>
<td>Argument nominalization: (Adnominal) Modification-use</td>
</tr>
<tr>
<td>g. Adverbial clause,</td>
<td>Event nominalization: (Adverbial) Modification-use</td>
</tr>
<tr>
<td>Converb, Clause-chain</td>
<td></td>
</tr>
<tr>
<td>h. Headless/Free genitive</td>
<td>Nominal-based nominalization: NP-use</td>
</tr>
<tr>
<td>i. Genitive/Possessive</td>
<td>Nominal-based nominalization: (Adnominal) Modification-use</td>
</tr>
<tr>
<td>j. (Numeral) classifier</td>
<td>Nominal/Numeral/Verbal-based nominalization</td>
</tr>
</tbody>
</table>

9.1 Implications

The crosslinguistic study of nominalizations and their roles in grammar presented above has some far-reaching implications for both descriptive practice and theoretical debates. Indeed, the philosopher of language Zeno Vendler, who also worked on nominalization, noted that “the grammar of nominalizations is a centrally important part of linguistic theory” (Vendler 1967: 125). Of the various implications to many parts of grammar, we limit our discussions here to those pertaining to the understanding and analysis of relative clauses because of the high impact that this research topic has had in the field over the past fifty years.

We start with the widely received definitions of relative clauses and their evaluations in the light of our analysis of nominalization in this paper. The following two definitions perhaps represent the general understanding in the field of what relative clause constructions are:

a construction consisting of a nominal or common noun phrase …which may be empty …and a subordinate clause interpreted attributively modifying the nominal. The nominal is called the head and the subordinate clause the RC. The
attributive relation between head and RC is such that the head is involved in what
is stated in the clause. (Lehmann (1986: 664))

A relative clause (RC) is a subordinate clause which delimits the reference of an
NP by specifying the role of the referent of that NP in the situation described by
the RC. (Andrews (2007: 206))

Both these definitions are based on the traditional clause-based analysis of
English-type relative clauses that recognizes an involvement of so-called relative
pronouns that link the head noun with an argument position in the RC structure
(our grammatical argument nominalization), as if the denotation/referent of the
head noun is directly represented in the RC structure (see (5-86)). Such descrip-
tions, however, fail to capture the notion of “restrictive” modification properly,
which involves two independent sets of entity denotations. If the head NP is core-
ferential with an NP in the relative clauses as in the traditional analysis, we are
dealing with a single referent; and the notion of restrictive modification does not
fall out from such an analysis. Our analysis of RC constructions does not recog-
nize as essential the English-type relative pronouns, which are not found in most
languages outside the Indo-European sphere. More importantly it is not really the
case that the denotation/referent of the head noun “is involved in what is stated
in the clause” or that their role is specified “in the situation described by the RC”.
Aside from the point that so-called RCs do not “state” or “describe” like clauses
and sentences, the denotations of the two nominals involved in RC constructions,
a head noun and a modifying argument nominalization, actually denote two dis-
tinct sets of denotation. The only and important requirement for the restrictive
RC construction is that the denotation of the argument nominalization be able to
specify a subset of that of the head noun, as in the manner shown in Figure 3. Our
analysis not only translates straightforwardly to the Formal Semantics treatment
(see page 84), but also captures in a very simple manner the intent behind the defi-
nition of the restrictive RC construction by Keenan and Comrie below, which suf-
fers from the fact that there is no evidence that a truth-bearing sentence underlies
an RC. Argument nominalizations qua RCs do not assert; they only presuppose.

[an RC construction] specifies a set of objects (perhaps a one-member set) in
two steps: a larger set is specified …and then restricted to some subset of which a
certain sentence, the restricting sentence, is true. The domain of relativization is
expressed in surface structure by the head NP, and the restricting sentence by the
restricting clause, which may look more or less like a surface sentence depending
on the language. (Keenan & Comrie 1977: 63-64)

Let us now turn to the single most influential paper on the topic of relative clauses,
namely Keenan & Comrie (1977). Our new analysis first calls for a reassessment of
the NP-Accessibility Hierarchy that plays the central role in the Keenan-Comrie analysis of RCs. Our analysis suggests that grammatical relations actually have nothing to do with relative clause formation per se, which is viewed as bringing together a grammatical argument nominalization and a head noun to form a larger NP constituent without ever asking whether what is being relativized on is Subject, Object, or Oblique. A so-called subject relative clause is simply a modification of a noun by a subject nominalization, and a so-called object RC is no more than bringing an object nominalization and a head noun together under the modification function. Under the proposed analysis of RC constructions, there is no process involved that “accesses” an argument position, as in the traditional generative analysis, which creates a gap in an argument position as part of the relativization process. In our analysis, a gap in the modifying structure of an RC construction is a property of an argument nominalization.

This does not invalidate a hierarchy of grammatical relations like the one posited by Keenan and Comrie. Indeed, such a hierarchy is plausible for argument nominalizations, since some languages, such as Yup’ik, allow only argument nominalizations of the absolutive argument. Apparently some dialects of K’iche’ are like Yup’ik, while other dialects allow argument nominalizations pointing to both absolutive and ergative roles, as shown in the examples cited in this paper (see Larsen & Norman 1979). Those Austronesian languages (many Formosan and Philippine languages as well as Malagasy) maintaining the proto-Austronesian four-way focus contrast allow argument nominalizations of various types, such as subject nominalization, object nominalization, locative, and beneficiary (see the Malagasy-German comparison in the following discussion). Those that have reduced the focus contrast to two (AF and PF), as in many languages of Indonesia (Bahasa Indonesia, Javanese, Balinese, Sasak, etc.), allow only subject and object nominalizations; obliques must be first made applicative objects before they can be the target of argument nominalization. The English and German gerundive argument nominalization applies only to subjects; e.g. the man [Ø holding a book in hand], *the book [the man holding Ø in hand]. Our point is that while argument nominalizations refer to grammatical relations, the relativization process itself does not, contrary to the claim advanced by Keenan and Comrie. Indeed, the relevance of relational hierarchies makes much more sense in the metonymy-based analysis of nominalization than the clause-based analysis of relativization. Since the absolutive/ergative and the subject/object arguments, as central argument types, code most salient event protagonists intimately associated with a wide range of event types, they are the easiest to evoke metonymically.

Our new analysis of RCs also has significant implications for the description of RCs in individual languages. In particular, it shows that those Austronesian languages (e.g. Malagasy and Tagalog), which are claimed to obey the subject-only
constraint on relativization actually relativize on any argument as does German, which is said to relativize down to the genitive position in the Accessibility Hierarchy. Keenan and Comrie demonstrate that relativization on a subject (9-1b) is possible in an Actor focus construction, but an object in such a construction cannot be relativized (9-1c). For an object to be relativized, it must be made into a subject by turning an AF construction to a Patient focus construction (9-2a).

(9-1) Malagasy AF construction
a. n-i-kapoka ilay alika t-amin y hazokely ilay lehilahi
   PST-AF-hit DEF dog PST-with stick DEF man
   ‘The man hit the dog with a stick.’

b. n-a-hita ilay lehilahy (izay) [n-i-kapoka ilay alika t-amin y]
   PST-AF-see DEF man NMLZR PST-AF-hit DEF dog PST-with
   hazokely Ø] aho stick 1sg (AF-NMLZR + SUB NMLZ)
   ‘I saw the man [who Ø hit the dog with a stick].’

c. *n-a-hita ilay alika (izay) [n-i-kapoka Ø t-amin y hazokely]
   PST-AF-see DEF dog NMLZR PST-AF-hit PST-with stick
   ilay lehilahi] aho
   DEF man 1sg
   ‘I saw the dog [that the man hit Ø with a stick].’

(9-2) Malagasy PF construction
a. no-kapoh-in ilay lehilahy t-amin y hazokely ilay alika
   PST-hit-PF-DEF man PST-with stick DEF dog
   ‘The man hit the dog with a stick.’

b. n-a-hita ilay alika (izay) [no-kapoh-in ilay lehilahy t-amin y]
   PST-AF-see DEF dog NMLZR PST-hit-PF-DEF man PST-with
   hazokely Ø] aho stick 1sg (PF-NMLZR + OBJ NMLZ)
   ‘I saw the dog that Ø was hit by the man with a stick.’

Assuming PF and other non-AF constructions to be passive, Keenan and Comrie conclude that only subjects can be relativized on in Malagasy, instantiating a language in which the subject-only constraint on relativization obtains (Keenan & Comrie 1977; Comrie & Keenan 1979). As it turns out, the real reason why (9-1c) is ungrammatical is not because what has been relativized on (the gap position) is object position, but because the construction has an incompatible combination of AF marking and object nominalization. Recall from the earlier discussion on another focusing Austronesian language, Mayrinax Atayal (see (3-12)), that focus marking in Austronesian has a role-marking function for argument nominalizations, where AF marking in the verb marks a subject nominalization, PF marking
an object nominalization, LF marking a locative nominalization, and CF marking a beneficiary or an instrumental nominalization. AF marking, therefore, can combine only with a subject nominalization, as in (9-1b), and cannot combine with an object nominalization, as in (9-1c). An object nominalization must be marked by the PF marker in the verb, as in (9-2b),\textsuperscript{101} not by the AF marker as in (9-1c).

Languages with role-marking nominalizers all behave this way, such that a subject/agent nominalizer must mark a subject nominalization, an object/patient nominalizer an object nominalization, and so forth. Indeed, the Malagasy pattern is paralleled by German, which also has role-marking nominalizers, similar to AF and PF markers in focusing Austronesian languages. Observe:

(9-3) German subject nominalization
      ART boy sees ART dog
      ‘The boy sees the dog.’
   b. [der Junge [der [Ø den Hund sieht]]]
      ART boy SUB.NMLZR ART dog sees (SUB.NMLZR + SUB NMLZ)
      ‘the boy who sees the dog’
   c. *[der Hund [der [der Junge sieht Ø]]]
      ART dog SUB.NMLZR ART boy sees (SUB.NMLZR + DO NMLZ)
      ‘the dog that the boy sees’
   d. [der Hund [den [der Junge sieht Ø]]]
      ART dog OBJ.NMLZR ART boy sees (DO.NMLZR + DO NMLZ)
      ‘the dog that the boy sees’

The reason that (9-3c) is bad is not because German cannot relativize on an object, but because it has the incompatible combination of a subject nominalizer and an object nominalization, as in (9-1c) for Malagasy. The phrase in (9-3d) is grammatical because the object nominalizer marks an object nominalization, just like the Malagasy form (9-2b). We see an exact parallelism between Malagasy and German. Indeed, Malagasy can relativize on any argument and adjunct that German can as long as the marking pattern is consistent. Just to drive the point home, another parallelism between the two languages using an oblique nominalization and its use as a modifier (relative clause) is shown below.

\textsuperscript{101} Like many other languages, e.g. Mongolian, Turkish, Japanese, Yaqui, and Quechua, object nominalizations in focusing Austronesian languages may have an agent in the genitive form.
(9-4) German oblique/source nominalization
   a. NP-use
      \[\text{Ich treffe den \{\text{vom \ dem ich das Buch bekommen habe}\}_{\text{NMLZ}}\}_{\text{NP}}\]
      I meet ART from IO.NMLZ I the book receive.PP have
      ‘I meet the one from whom I received the book.’
   b. Modification-use
      \[\text{Ich treffe den \{\text{Mann vom \ dem ich das Buch bekommen habe}\}_{\text{NMLZ}}\}_{\text{NP}}\]
      have
      ‘I meet the man from whom I received the book.’

(9-5) Malagasy oblique/source nominalization
   a. NP-use
      \[\text{Ho hita-ko ilay [n-\text{indrama-ko} (an’ilay/ilay) boky]}\]
      FUT see-1SG GEN the PST-borrow.CF-1SG GEN (the/the) book
      ‘I will see the one from whom I borrowed the book.’
   b. Modification-use
      \[\text{Ho hita-ko ilay lehilahy [n-\text{indrama-ko} (an’ilay/ilay) boky]}\]
      book
      ‘I’ll see the man from whom I borrowed the book.’

The parallelism between German and Malagasy is again clear—(9-4b) and (9-5b)
are grammatical RC constructions because the nominalizers mark correct nominal-
ization structures in both cases. If the nominalizers in these examples were
the subject nominalization marker \textit{der} for German or the subject nominalizing
AF form for Malagasy, both would be ungrammatical. As long as nominalizers
and nominalization structures are kept consistent, both languages can nominalize
any argument down to the genitive in the relational hierarchy, and the resulting
nominalizations can be used as modifiers (relative clauses). We can observe the
same thing in English, which has an object nominalizer that uniquely marks hu-
man object nominalizations, and which, therefore, cannot combine with a subject
nominalization, as in (9-6b) below.

(9-6) English object and subject nominalizations
   a. \textit{Marry [who/whom [you love \Ø]]} (object nominalization)
   a’. \textit{Marry a man [who/whom [you love \Ø]]}
   b. \textit{Marry [who/*whom [\Ø loves you]]} (subject nominalization)
   b’. \textit{Marry a man [who/*whom [\Ø loves you]]}
As is clear from the above, the proposed nominalization-based analysis of relative clause constructions yields a very different result from the traditional clause/sentence-based analysis by Keenan and Comrie (1977), Comrie and Keenan (1979) and others. This is true of all focusing Austronesian languages such as Atayal and other Formosan languages, Tagalog and other Philippine languages, Malay/Indonesian, Sasak, Sumbawa, and others (see Shibatani (2008) on Sasak and Sumbawa).

There are many other important theoretical and descriptive implications engendered by the proposed analysis of nominalization, but let us close this already lengthy paper by touching on the problems of the power of a grammatical theory. The analysis of so-called relative clauses and complement clauses as nominalizations rather than as clauses allows a much tighter theoretical framework for syntactic analysis; namely only structures of equal or lower rank can be embedded within a given structure. Current theories, which consider relative clauses and complement clauses as clauses, allow clauses to be embedded under a structure lower in rank such as NP and VP. Such theories allowing any type of embedding are too powerful and hence are weak theories.

As these brief comments suggest, the new analysis of nominalizations proposed in this paper opens up many interesting new developments in both descriptive and theoretical arenas.

Acknowledgements

This chapter is a vastly expanded version of an earlier paper, Shibatani (2018a), which contrasts Japanese nominalizations with those of the world’s languages. The present paper, on the other hand, draws data from the Americas and elsewhere and offers more detailed discussions on both conceptual and empirical issues surrounding nominalization phenomena. I owe a great deal of indebtedness to many scholars who have helped me gather and interpret the nominalization data contained in this paper, but especially to Akio Ogawa (German), Roberto Zariquiey (Spanish), Cristina Messineo (Toba), Jane Barnes (Tuyuca), Doris Payne (Yagua), Marilia Ferreira (Parkatêjê and Portuguese), Telma Can Pixabaj and Mareike Sattler (K’iche’), Kingkarn Thepkanjana and Shoichi Iwasaki (Thai), Haowen Jiang (Mandarin Chinese, Cantonese), Sung-Yeo Chung (Korean), Stephen Matthews and Virginia Yap (Cantonese), Nerida Jarkey, Martha Ratliff, and Elizabeth Riddle (White Hmong), Kazuyuki Kiryu (Newar), Prashant Pardeshi (Marathi), Miki Nishioka (Hindi), and Albert Alvarez (Yaqui). I am also grateful to Francesc Queixalós, David Fleck, and Roberto Zariquiey for many useful discussions on nominalization. I am also grateful to David Fleck and Haowen Jiang, who read an earlier version carefully and helped spot typos and make stylistic improvements. The research and preparation of the work reported here was in part supported by the International Joint Research Promotion Program of Osaka University (PI: Sung-Yeo Chung) and the project “Noun Modifying Expressions” (PI: Prashant Pardeshi) of the National Institute for Japanese Language and Linguistics (NINJAL).
### List of less commonly used abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF</td>
<td>actor focus</td>
</tr>
<tr>
<td>ANP</td>
<td>animate plural</td>
</tr>
<tr>
<td>BUT</td>
<td>contraexpectation</td>
</tr>
<tr>
<td>C</td>
<td>time reference 'before', 'completed'</td>
</tr>
<tr>
<td>CA</td>
<td>causative</td>
</tr>
<tr>
<td>CF</td>
<td>circumstantial focus</td>
</tr>
<tr>
<td>CL</td>
<td>clitic</td>
</tr>
<tr>
<td>CLF</td>
<td>classifier</td>
</tr>
<tr>
<td>CM</td>
<td>concatenated form</td>
</tr>
<tr>
<td>COLL</td>
<td>collective</td>
</tr>
<tr>
<td>CON</td>
<td>conjunctive particle</td>
</tr>
<tr>
<td>CONJ</td>
<td>conjecture</td>
</tr>
<tr>
<td>DE</td>
<td>demonstrative pronoun</td>
</tr>
<tr>
<td>DIR</td>
<td>directional</td>
</tr>
<tr>
<td>DL</td>
<td>dual</td>
</tr>
<tr>
<td>DM</td>
<td>'uh', preform</td>
</tr>
<tr>
<td>DT</td>
<td>associated motion marker</td>
</tr>
<tr>
<td>EVD</td>
<td>evidential</td>
</tr>
<tr>
<td>EX</td>
<td>existential</td>
</tr>
<tr>
<td>F</td>
<td>feminine</td>
</tr>
<tr>
<td>FGR</td>
<td>falling tone grade</td>
</tr>
<tr>
<td>FSG</td>
<td>feminine singular</td>
</tr>
<tr>
<td>FV</td>
<td>final vowel</td>
</tr>
<tr>
<td>GER</td>
<td>gerundive</td>
</tr>
<tr>
<td>HAB</td>
<td>habitual</td>
</tr>
<tr>
<td>HON</td>
<td>honorific</td>
</tr>
<tr>
<td>INC</td>
<td>inceptive aspect marker</td>
</tr>
<tr>
<td>IR</td>
<td>irealis</td>
</tr>
<tr>
<td>LF</td>
<td>locative focus</td>
</tr>
<tr>
<td>LGR</td>
<td>lengthened grade</td>
</tr>
<tr>
<td>LIN</td>
<td>linker</td>
</tr>
<tr>
<td>MM</td>
<td>middle marker</td>
</tr>
<tr>
<td>MOT</td>
<td>motion</td>
</tr>
<tr>
<td>MSG</td>
<td>masculine singular</td>
</tr>
<tr>
<td>MPL</td>
<td>masculine plural</td>
</tr>
<tr>
<td>NAR</td>
<td>narrative register</td>
</tr>
<tr>
<td>ND</td>
<td>neutral disjunct</td>
</tr>
<tr>
<td>NPM</td>
<td>NP-use marker</td>
</tr>
<tr>
<td>NMLZ</td>
<td>nominalization</td>
</tr>
<tr>
<td>NMLZR</td>
<td>nominalizer</td>
</tr>
<tr>
<td>MOD</td>
<td>modal clitic</td>
</tr>
<tr>
<td>O</td>
<td>object</td>
</tr>
<tr>
<td>OBJ.An</td>
<td>animate object</td>
</tr>
<tr>
<td>OM</td>
<td>object marker</td>
</tr>
<tr>
<td>P</td>
<td>possessor</td>
</tr>
<tr>
<td>PF</td>
<td>patient focus/affirmative evidential</td>
</tr>
<tr>
<td>PFV</td>
<td>perfective</td>
</tr>
<tr>
<td>PR</td>
<td>progressive</td>
</tr>
<tr>
<td>PROX</td>
<td>proximate</td>
</tr>
<tr>
<td>PRT</td>
<td>particle</td>
</tr>
<tr>
<td>PS</td>
<td>optional past tense marking</td>
</tr>
<tr>
<td>REF</td>
<td>referential</td>
</tr>
<tr>
<td>REL</td>
<td>relative case (Yup’ik)</td>
</tr>
<tr>
<td>S</td>
<td>subject</td>
</tr>
<tr>
<td>SIN</td>
<td>single action</td>
</tr>
<tr>
<td>&lt;T&gt;</td>
<td>verb-terminating classifier</td>
</tr>
<tr>
<td>THM</td>
<td>theme (grammatical relation)</td>
</tr>
<tr>
<td>TOP</td>
<td>topic marker</td>
</tr>
<tr>
<td>TR</td>
<td>transitivizer</td>
</tr>
<tr>
<td>TRZ</td>
<td>transitivizer</td>
</tr>
<tr>
<td>VI</td>
<td>verbal stem marker</td>
</tr>
<tr>
<td>~</td>
<td>non-</td>
</tr>
</tbody>
</table>

### References


https://doi.org/10.1016/B0-08-044854-2/01111-1


Campbell, Akua Asantewaa. 2013. Grammatical nominalizations in Akan and Gã. Ms, Department of Linguistics, Rice University.


Chapter 2. What is nominalization? Towards the theoretical foundations of nominalization

https://doi.org/10.1017/CBO9780511778322


https://doi.org/10.1524/stuf.2004.57.23.247


https://doi.org/10.4324/9780203340769


https://doi.org/10.1075/tsl.76.09hor


https://doi.org/10.1515/9783110864823


Tamura, Yuki-Shige. 2017. Grammaticalization and nominalization in Central Alaskan Yup’ik. Presentation at the Nominalization Festival III, Osaka University, 8 July.


