Nominalization in Soqotri, a South Arabian language of Yemen

Masayoshi Shibatani (Rice University) and Khaled Awadh Bin Makhashen (Hadhramout University and Universiti Sains, Malaysia)

1. Introduction

In his recent article Deutscher (2009: 199) suggests that "nominalization is an unsung hero in the story of subordination". He goes on to say that "[t]he ability to derive a noun from a verb, that is, to reify a verbal predicate and present it as a nominal argument or modifier, is the core of subordination". While Deutscher (2009) believes that the relevant nominalization for his study is "derivational nominalization", which derives nouns from verbs, we find his characterization of nominalization as an unsung hero to be quite fitting not only in the study of subordination but also in other related phenomena, which we wish to characterize as the case of "grammatical nominalization". Unlike derivational or lexical nominalization of the *employ-er/employ-ee* type, which supplies the lexicon with new lexical items belonging to the noun class, grammatical nominalization yields nominal expressions but which are not nouns. We shall explicate the difference between lexical nominalization and grammatical nominalization below, from which various properties of grammatical nominalizations follow.

The issues addressed in this paper are those nominal expressions that have been discussed under the heading of "noun phrases without nouns" by Matthew Dryer in his recent review of noun phrase structure (Dryer 2007), where he surveys noun phrases without a nominal head such as adjectives functioning as noun phrases by themselves (see (1) below), possessor phrases without a noun head (see (2) below), and so-called headless relative clauses (see (3a) below):

(1) Nkore-Kiga (Dryer 2007:194 quoting Taylor 1985)
[omuto] aka-gamba na-anye
young 3SG-REM.PAST-speak with-me
'the young one spoke with me'

(2) Your car is nice, but *John's* is nicer.

(3) Miya (Dryer 2007:197 quoting Schuh 1998)
a. má ráɓaza (headless RC)
REL.FEM.SG wet
'the one (feminine, singular) that is wet'
b. kàba [má ráɓaza] (headed RC)
gown REL.FEM.SG. wet
'the gown that is wet'

While Dryer (2007) does not invoke the notion of grammatical nominalization for any of these constructions above, some recognize the involvement of nominalization in some of them. For example, Comrie & Thompson (2007: 378, 379) recognize what they regard as "a somewhat more rare function of nominalization: as a relative clause modifying a head noun…" and make the puzzling "claim that in certain languages relativization is indistinct from nominalization", as if relative clauses and nominalizations were two separate constructions that happen to be structurally indistinct in some languages. A less misleading way to characterize the situation that Comrie & Thompson (2007) observe is to make a clearer distinction between form and function, and to consider that relativization and nominalization are two different functions – modification for the former and referential for the latter –, and then to recognize that one and the same form plays both functions in some languages.

Besides Dryer a countless number of grammarians characterize forms such as Miya example (3a) above as headless relative clauses as if they were derivatives of (headed) relative clause constructions. Indeed, the most popular analysis of such forms is a deletion analysis, where the relevant forms (e.g. 3a) are derived from headed relative clauses (e.g. 3b) via deletion of the head nominal. We consider this kind of analysis misguided and instead regard forms such as (3a) as simply nominalizations. We consider relative clauses of the type seen in (3b) above to be not independent constructions which happen to be structurally indistinct from nominalizations but that they are nominalizations themselves that play a modification function. This position has recently been expounded in Shibatani (2009), where it is shown that a nominalization functioning as a relative clause is not at all rare, contrary to the suggestion made by Comrie & Thompson (2007; see above), and that it is widely seen around the

globe, and even in familiar European languages. Our attempt here is to show that nominalizations are involved not only in relative clauses (of both headed and headless types) but also in other types of noun phrases surveyed by Dryer (2007). The relevant data derive from the endangered language Soqotri of Yemen, which displays a wide array of nominalization phenomena, ranging from lexical to grammatical nominalization, and from argument nominalization to different types of event nominalization. Our primary focus in this paper is the phenomenon of grammatical argument nominalization.

2. Soqotri

The Soqotri language is one of six pre-Islamic languages that form a group called Modern South Arabian languages. This group, which includes Soqotri, Mehri, Jibbali, Bathari, Harsusi and Hobyot, is spoken in parts of Oman particularly in Dhofar and south east of Yemen in Mahrah Governorate and Soqotra Island. Modern South Arabian languages belong to the southern branch of the western Semitic languages. The other Semitic languages grouped in this branch are the Semitic languages of Ethiopia and the extinct inscription languages, which are also called Epigraphic Old South Arabian languages. Soqotri is the native language of the Island of Soqotra located in the east of Aden Gulf about 300km south of the Arabian Peninsula. It is also spoken in two small nearby islands called the Island of Abdul kuri and the Island of Samha. The number of the Soqotri speakers is estimated to be around 50,000. The island inhabitants depend on fishing, growing palm dates trees or rearing camels, cows and goats. The Soqotri language and culture is under a great influence of the dominant Arabic language and culture, and it is regarded as an endangered language (Naumkin 1998, Simeone-Senelle 2003). Many Arabic-speaking Yemenis have settled in the Soqotri territory permanently. Arabic has become the official language in the island, and it is used in the Soqotri schools as medium of instruction. The Soqotri students are prohibited from using their mother tongue while they are at school. Any Soqotran seeking employment must master Arabic before landing a reasonable job. The young Soqotrans, fluent in Arabic, prefer it to their mother tongue, which they learn imperfectly, mixing many Arabic words in it. They cannot recite or understand any piece of Soqotri oral literature.

Soqotri is primarily a head-initial language with VS/SV and VSO/SVO word orders seen equally frequently. It has prepositions, and adjectives and relative clauses as well as genitives (typically with a full possessor noun) follow the head noun they modify, while demonstratives, numbers, and genitives (typically with a pronominal possessor) precede the head noun. The modifiers agree in number and gender with the head noun, as shown below:

(4) dɛ-**ʃi** ʁəɬɛːt**i** ʕɛlɛːt**i** dħəh dɛ-**ʃi** ɬədɛːħ-**tɛh**
DEM-DU.F woman-DU.F tall-DU.F here REL.DU.F come-3DU.F.PER
nɑːʕɛ**h** dɛhɒ χɛlɒ**ti**
now my aunt-DU.F
'These two tall women who came just now are my aunts.'

3. Relative clauses and other types of noun modification

Of the several types of noun modification in Soqotri, the patterns we are interested in here are relativization, adjectival modification and genitive constructions, all of which share what we claim to be the basic nominalization mechanism of the language. As seen in (4) above, Soqotri relative clauses are marked by what is identified as the "relative particle" by Simeone-Senelle (1997). It inflects for number and gender in accordance with these categories of the head noun.

(5) a. ʁəɬ dɛ jə-ʕɑːɮɛn birhɛ ɬədəħ
man REL-SG.M 3SG.M.IMP-love children come-3SG.M.PER
'A man who loves children came.'
b. ʁəɬ**i** dɛ**ki** jəʕəɮɛːnɛ**h** birhɛ ɬədɛːħɛ**h**
man-DU.M REL DU.M 3DU.M.IMP-love children come-3DU.M.PER
'Two men who love children came.'
c. ʁ**ɪɒ**ɬ **l**ɛ **jə**-ʕɛːɮɛn birhɛ ɬəd**uː**ħ
men REL-PL 3PL.M.IMP-love children come -3PL.M.PER
'Men who love children came.'

(6) a. ʃiːn-ək ʁɑːɬɛh **dɛʃ** **tə**ʕɑːɮɛn birhɛ
saw-1SG.M.PER woman REL SG.F 3SG.F.IMP-love children
'I saw a woman who loves children.'

b. ʃiːn-ək ʁɑːɟiː**ti** **dɛʃi** təʕɛːɻnɛ**h** birhɛ
saw-1SG.M.PER woman-DU REL.DU.F 3DU.F.IMP-love children
'I saw two women who love children.'

c. ʃiːn-ək ʁəɟɛː**tɛn** **lɛ** təʕɑːɻənɛ**n** birhɛ
saw-1SG.M.PER women REL.PL 3SG.F.IMP-love children
'I saw women who love children.'

Keenan (1985:149) notes that relative pronouns are "typically the same as, or morphologically related to, the demonstrative pronouns". This is true for the Soqotri "relative particles", which are identical with the demonstrative pronouns. The Soqotri demonstratives combine demonstrative pronouns and locative adverbials *dħəh* (SG/DUAL) and *lħəh* (PL) 'here' for proximal and *dbɒ k* (SG/DUAL) and *lbɒk* (PL) 'there'. Demonstrative pronouns may also be followed by these locative adverbials.

Number	Masculine		Feminine	
Singular	dɛ…dħəh	'this'	dɛʃ…dħəh	'this'
Dual	dɛki…dħəh	'these'	dɛʃi…dħəh	'these two'
Plural	lɛ…lħəh	'these'	lɛ…lħəh	'these'

Table 1: Proximal demonstratives

Number	Masculine		Feminine	
Singular	dɛ. . . dbɒ k	'that'	dɛʃ...dbɒk'	'that'
Dual	dɛki…dbɒk'	'those'	dɛʃi…dbɒk'	'those two'
Plural	lɛ…lbɒk	'those'	lɛ…lbɒk	'those'

Table 2: Distal demonstratives

(7) a. dɛ (dħəh) ø dɛhɒ mɒkʃəm
DEM here COP my child
'This (here) is my child.'

b. ʕiːk lɜ-tɟɛr lɛ lɛbɒk'
want-3SG.IMP 1SG.SUB-buy DEM.PL there.PL
'I want to buy those (FEM/MAS).'

c. dɛ mɒkʃəm dħəh ∅ dɛhɒ mɒkʃəm
DEM boy here COP my child
'This boy is my child.'

d. ʕiːk lɜ-tɟɛr lɛ ʔɛlhɛːtɛn lbɒk'
want-3SG.IMP 1SG.SUB-buy DEM.PL cow-PL.F there.PL
'I want to buy those cows'

What can be characterized as headless relative clauses in Soqotri show formal resemblance to headed relative clauses as in other languages possessing these two types of constructions, as comparison of the following with the corresponding relative clauses above indicates.

(8) a. dɛ jə-ʕɑːɮɛn birhɛ ɟədəħ
REL.SG.M 3SG.M.IMP-love children come-3SG.M.PER
'The one who loves children came.' (cf. 5a)

b. dɛ**ki** jəʕəɮɛːnɛ**h** birhɛ ɟədɛːħɛ**h**
REL.DU.M 3DU.M.IMP-love children come-3DU.M.PER
'The ones who love children came.' (cf. 5b)

c. lɛ **jə**-ʕɛːɮɛ**n** birhɛ ɟəd**uː**ħ
REL.PL 3PL.M.IMP-love children come-3PL.M.PER
'The ones who love children came.' (cf. 5c)

(9) a. ʃiːn-ək **dɛʃ** **tə**ʕɑːɮ ɛn birhɛ
saw-1SG.M.PER REL.SG.F 3SG.F.IMP-love children
'I saw the one who loves children.'(cf. 6a)

b. ʃiːn-ək **dɛʃi** **tə**ʕɛːɮ nɛh birhɛ
saw-1SG.M.PER REL.DU.F 3DU.F.IMP-love children
'I saw the ones who love children.' (cf. 6b)

c. ʃiːn-ək **lɛ** təʕɑːɮ ənɛn birhɛ
saw-1SG.M.PER REL.PL.F 3SG.F.IMP-love children
'I saw the ones who love children.' (cf. 6c)

The reason we enclosed the term "relative particles" in the quotation marks above is that the forms *dɛ, dɛʃ*, etc. that introduce relative clauses are used in other constructions besides relative clauses and the so-called headless relative clauses seen above. They are used in adjectival attributive constructions and genitive, or more broadly noun modifier, constructions.

Like English, adjectives in Soqotri can modify nouns directly or in a form similar to relative clauses, as below:

(10) a. ʁɑːɟɛh ʃkɛːrɛh
woman beautiful
'a beautiful woman' or 'The woman is beautiful.'
b. ʁɑːɟɛh **dɛʃ** ʃkɛːrɛh
woman REL.SG.F beautiful
'a beautiful woman' or 'the woman who is beautiful'

Apparently (10b) above allows both attributive and restrictive relative readings. Adjectives differ from verbs in that while the former can modify a noun directly, as in (9a), the latter cannot. For example, (11a) below cannot mean 'a running woman'.

(11) a. ʁɑːɟɛh təʃɒʕ
woman 3SG.F.IMP-run
'The woman runs.'
b. ʁɑːɟɛh **dɛʃ** təʃɒʕ
woman REL.SG.F 3SG.F.IMP- run
'the woman who runs'

Genitive constructions in Soqotri come in three types. In expressing kinship relations, the kin terms are simply marked by bound object pronouns (12). The other two involve the marking relevant to our main discussion. One of them involves the "possessive particle" *dɛ* and free pronouns or nouns, as in (13), and the other the prefix *m-* and bound object pronouns, as in (14).

(12) a. ʔɛħ **-i**
brother-1SG.M.OBJ.SUF
'my brother
b. ʔɛħ **-ək**
brother-SG.M.OBJ.SUF
'your brother'
c. ʔɛħ **- əs**
brother-3.SG.F.OBJ.SUF
'her brother'

(13) a. **dɛ**hɒ fɑːnɛ
my face
'my face'
b. fɑːnɛ **dɛ** ʁəɟ
face POSS. man
'the man's face'
c. fɛrhim **dɛ** ʕəli
daughter POSS Ali
'Ali's daughter'

(14) a. **m**ənhɛ fɑːnɛ
my face
'my face'

b. **m**ək fɑːnɛ
your face
'your face'

c. **m**ɛh fɑːnɛ
his face
'his face'

The genitive constructions involving the "possessive particle" seen in (13) are in fact general noun modifier constructions used in the modification of a noun by another noun, as in the following forms:

(15) a. mdərrɛs **dɛ** mədrəsɛh
teacher POSS school
'school teacher'

b. ktɑːb **dɛ** ħsɑːb
book POSS mathematics
'mathematics book'

c. tɛh **dɛ** tɛːtɛn
meat POSS sheep
'sheep meat'

d. mɛssə **dɛ** χtɛːr
rain POSS night
'night rain'

While the "possessive particle" involved here is obviously related to the "relative particles" seen earlier, they behave somewhat differently. Namely, the "possessive particle", unlike the "relative particle", does not inflect according to the head noun. The form *dɛ* is the masculine singular form, which is also used as a default form when the gender and number of the referent is indeterminate (see (23) below). The reason why the "possessive particle" does not inflect is not obvious at this point. Besides the formal resemblance between the "relative" and the "possessive" particle, there is some indication that the relevant constructions involve the same mechanism. Namely, the *m-* possessive prefix seen in (14) appears to be the same prefix used in lexical argument nominalization that yields forms like the following (see below for more examples):

(16) a. **m** ʃɛʕrɛk (N) 'fisherman'
b. **m**s'əlli (N) 'prayer/person who prays'
c. **m**ʃɛħlɛf (N) 'sharpener'
d. **m**ɒk'dɛf (N) 'oar'

This sharing of the marker *m-* by both argument nominalization and genitive constructions is indicative of the fact that nominalization is involved in the relevant genitive construction. Along the same vein, we can safely hypothesize that both relative clause constructions and the other

modification patterns observed above involve nominalization resulting in *dɛ*-marking seen in all these noun modification constructions.

This hypothesis above is bolstered by the similarity between relative clauses and the adjectival and noun modifier (including the genitive) constructions that goes beyond the sharing of the *dɛ*-marking. Just as relative clauses have "headless relative clause" counterparts, both adjectival and noun modifier constructions allow "headless" expressions, as seen below:

(17) "Headless adjectival modifiers"

a. ʁəʄ dɛ ħɑːr ʄədəħ
man REL.SG.M black.3SG.M came.3SG.M.PER
'A black man came.'

a'. dɛ ħɑːr ʄədəħ
REL.SG.M black.3SG.M came.3SG.M.PER
'A black (one) came.'

b. bɑːʕɛl ʁɑːʄɛh dɛ(ʃ) -ʃkɛːr-ɛh
marry.3SG.M.PER woman REL.SG.F-beautiful-3SG.F
'He married a beautiful girl.'

b'. bɑːʕɛl dɛ(ʃ) -ʃkɛːr-ɛh
marry.3SG.M.PER REL.SG.F-beautiful -3SG.F
'He married a beautiful one.'

(18) "Headless noun modifiers"

a. dɛʃ kʼɑːʕɛr dħə ∅ dɛhɒ kʼɑːʕɛr
DEM.SG.F house here COP my house
'This house is my house.'

a'. dɛʃ kʼɑːʕɛr ∅ dɛhɒ
DEM.SG.F house COP mine
'This house is mine.'

b. dɛʃ ∅ dɛhɒ fɛrhim w dɛʃ dbɒkʼ ∅ dɛh fɛrhim
DEM.SG.F COP my daughter and DEM there COP his daughter
'This is my daughter and that is his daughter.'

b'. dɛʃ ∅ dɛhɒ fɛrhim w dɛʃ dbɒkʼ ∅ dɛh
DEM.SG.F COP my daughter and DEM there COP his
'This is my daughter and that is his.'

All this consideration suggests reconsidering the treatment of *dɛ*-marking above, where it has been sometimes glossed as REL as if it were a relativizer, following Simeone-Senelle's (2003) labeling of it as the "relative particle" or as POSS, following her labeling of it as the "possessive particle", as if it were a form separate from what has been glossed as REL. A more serious problem arises in identifying the function of *dɛ*-marking in (8-9) and (17-18), where it occurs in constructions without an obvious head.

4. Are these all headless constructions?

It is clear why constructions like (8) and (9) are given an oxymoronic appellation of "headless relative clauses" or "free relatives" by those who have been concerned with the typology of relative clauses such as Keenan (1985), Dryer (2005, 2007), Andrews (2007) and Huang (2008); across languages these constructions are (nearly) identical in form with the ones used to modify a nominal head, i.e., the regular relative clauses (e.g., (5) and (6) above). The basic question never raised by those treating constructions like (8) and (9) as headless relative clauses is whether the relevant constructions fit the functional definitions given to relative clause constructions. Consider the following definitions offered for relative clause constructions:

(19) "[a relative clause 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." (Keenan & Comrie 1977: 63-64)

> "a construction consisting of a nominal…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)

It is clear that the constructions in (8) and (9) and similar ones in other languages do not really function to specify "a subset of objects in two steps" as stated in the definition by Keenan and Comrie or to modify a nominal element, as in Lehmann's definition. Indeed, we claim that the

constructions in (8) and (9) are not modifiers at all. Rather they are nominal expressions in their own right, i.e., nominalizations, that refer to an entity characterized in terms of the event in which it has crucial relevance, e.g., being involved as a subject or an object (see next section on the nature of the argument nominalizations alluded to here).

The traditional way of fitting so-called headless relative clauses to the mold of the standard relative clause construction is in terms of a deletion analysis, in which these forms are derived from headed relative clauses by deleting the head nominals (see Adams (1972) on Ancient Greek, Weber (1983) on Quechua, Sneddon (1996) on Indonesian[1], Huang (2008) on Qiang, Wrona (2008) on Old Japanese, etc., etc.). We consider this approach to so-called headless relative clauses to be wrong-headed and totally unmotivated in some cases.

Treating the so-called headless relative clauses as a species of relative clauses is obviously due to the wrong perspective that the linguists such as those named above have had about these constructions, namely viewing them from the perspective of relative clauses, which contain them. For more than thirty years, relative clauses have been a center piece in both Generative Grammar, where they motivated syntactic rules of movement/extraction and deletion, and typological studies, where cross-linguistic patterns of relativization have been pursued from several different angles, including the most influential seminal work in this area by Keenan & Comrie (1977). We consider this situation similar to the effort of a Japanese person trying to characterize tuna fish in terms of *sashimi*, as if tuna fish were a derivative of *sashimi* rather than the other way around. Our position, as suggested above, is that the so-called headless relative clauses such as those seen in (8) and (9) are not relative clauses at all. They are nominalizations that are independent of relativization, but which can be used in the relativization context as a modifier of a nominal head. In many other languages, similar nominalizations are used in other contexts, such as noun and verb complements and other types of subordinate constructions, besides the relativization context, just as tuna is used to make other dishes besides *sashimi*. We also advance the same nominalization analysis for the other "headless constructions" examined in the preceding section.

---

[1] Sneddon (1996) at least recognizes what is derived from deleting the head nominal as a nominalization.

At the empirical front, there are several nominalizations that cannot be derived from headed relative clause constructions via deletion of the head. For example, several European languages have nominalizations that do not occur in the relativization context (see Shibatani (2009 for the relationship between nominalization and relativization in European languages and elsewhere). For example:

(20) Russian
a. (tot,) [kto vymyl ruki], mozhet nachatj jestj
(that) who 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.'

(21) German
a. Ich empfange, [wer (auch) morgen kommt]
I receive who (also) tomorrow comes
'I receive who(ever) comes tomorrow.'
b. *Ich empfange den Mann, [wer morgen kommt]
I receive ART man who tomorrow comes
'I receive the man who comes tomorrow.'

(22) Spanish
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.'

There are also cases of nominalization where there is no conceivable head since the relevant expressions refer to a non-specific or generic entity, as in the following Soqotri examples:

(23) a. dɛ [j-ʃɛmtɛl sɒk'ɒt'riəjh swə] jə-fɒz
NMZ.SG.M 3SG.M.IMP-speak Soqotri well 3SG.M.IMP-win
'The one who speaks Soqotri well will win (the contest).'

b. ʔə-ʃiːnə dɛ [jəɟɛdħɛn k'ərɛːrɛh]
3SG.M.IMP-see NMZ.SG.M 3SG.M.IMP-come tomorrow
'I will see who(ever) comes tomorrow.'

c. ħɛhɛ ɑːl j-ətə dɛ [nɛːɛh b ħɑːdɛb]
person not 3 SG.M.IMP-eat NMZ.SG.M fall-3SG.M.PER on earth
'One shouldn't eat what has fallen on the floor.'

A possible response to these criticisms against the deletion analysis is to posit an empty pronominal head such as the arbitrary PRO, which has no phonetic substance but which refers to an arbitrary entity as in (23), or the anaphoric pro referring to some antecedent for the form seen in (24) below, for example. Both deletion and empty pronoun analysis appear motivated because the kind of constructions under consideration typically occur in the context where there in an "antecedent", as in the following example, where the "antecedent" is understood, or where they refer to a generic or an arbitrary entity as in (23) above.

(24) ʕəli jə-k'ɔːrɛ ktɑːb dɛ dɛh bijɒh ʔɛndək'
Ali 3SG.M.IMP-read book NMZ.SG.M his father 3SG.M.PER-give
ɛh tɛh w ʔəh ʔək'ɔːrɛ dɛ [dɛhɒ
him (Ali) it (book) and I read NMZ.SG.M my
məʕəllɛm ʔɛndək' ɛnhɛ tɛh]
teacher 3SG.M.PER-give me it
'Ali is reading the book which his father gave him, and I am reading the one that my teacher gave me.'

Similar deletion and empty pronoun analyses can be advanced for those "headless" constructions of other noun modification constructions such as adjectival and noun modifier constructions seen above. For example, Matisoff (1973), while arguing against the deletion analysis of the "headless relative clauses", subscribes to the deletion analysis of the "headless" genitive constructions of the Tibeto-Burman language Lahu. Advancing a deletion or an empty pronominal analysis for these "headless" constructions because they typically occur in the environments noted above is like calling whale fish because they live in the ocean. The crux of the problem is arriving at a proper understanding of nominalizations that would explain why they typically occur in those contexts, just as a proper understanding of whale would explain why they live in the ocean

despite the fact they are not fish. Indeed, the past works (e.g., Noonan (1997) and Wrona (2008)) on nominalizations fail to explicate what nominalization really is, concentrating more on their distribution pattern and historical development. Our position is that once we come to a proper understanding of the relevant nominalizations, their distributional and historical development patterns also yield to a natural understanding. The nominalization analysis of the relevant "headless" constructions advanced here regards these constructions to have no other head, which is deleted or which takes empty pronominal forms. Instead, the nominalizations themselves are heads in these constructions.

5. What is nominalization?

The essence of nominalization is creation of a referring expression, hence its sharing this essential nominal function with nouns, which refers to a state of affairs characterized by an event denoted by the clause (event nominalization), to an entity characterized in terms of the denoted event in which it has crucial relevance (argument nominalization), or to an entity having crucial relevance to the referent of a noun ("genitive" nominalization). There are important distinctions between what can be termed "lexical nominalization" and "grammatical nominalization". While both lexical and grammatical nominalization processes yield referring expressions and the both types of nominalization fill the syntactic function as arguments or predicate nominals just like simple nouns, the former creates nouns (e.g., *employ-er/employ-ee*) that belong to the noun class of the lexicon, and the latter creates nominal expressions that have no lexical status (e.g., the equivalents to the English expressions *(the one) who employ/(the one) whom someone employs*). Secondly, lexical nominalizations, like other simple nouns, identify or put a label on the referents and classify them as belonging to a particular class of entities denoted by the label, grammatical nominalizations lack identification or names. Thus, the English expression *what I bought yesterday* or its equivalents in other languages, for example, characterizes an entity referred to in terms of an event of my buying it yesterday, but its identity is not specified – it could be a book, a newspaper, a hamburger, an umbrella, or any other things that could be bought.

The properties and their distribution of grammatical argument nominalizations follow from this process of creating a variety of new re-

ferring expressions pertaining to a limitless range of everyday events that vary in their content. First, argument nominalization is associated with a presupposition that an event characterizing the entity referred to has taken place. The expression *what I bought yesterday* presupposes that I bought something yesterday. This property of argument nominalization carries over to its relativization function such that relative clauses represent presuppositions, as *the book which I bought yesterday* presupposes that I bought a book yesterday. Thus, neither in the nominalized form nor in its relative clause function, is there assertion of the state of affairs denoted in the presupposition even if the form may contain a finite verb form as in the English examples here. This is an important distinction between sentences and nominalized forms, and which allows the latter to be embedded or subordinated into a main assertive clause.

Since argument nominalizations are typically created for the nonce, they are often indefinite, and this accounts for the use of indefinite interrogative pronouns such as *what* and *who* as a place holder for the gap created by the nominalization process, as in a number of European languages. Since nominalizations are nominal entities, they often call for an article or other types of determiners just like regular nouns, as also in many European languages.[2] This is what we see in Soqotri, where nominalizations are marked by demonstratives. That is, what has been identified as "relative" and "possessive" particles by Simeone-Senelle (1997) arose from demonstratives marking nominalizations as nominal entities. While Soqotri demonstratives and definite articles used in other languages normally mark definite nominals, their use in marking nominalizations has resulted in the grammaticalization of them as nominalization markers such that they no longer mark only definite nominals. This is what we see in forms such as (23) above. Further indication of this is that even in relative clause constructions, some speakers do not inflect the nominalization marker and simply use the default (masculine) form *dɛ*. In other words, the best analysis of the so-called "relative" and "possessive" particles in Soqotri is to treat them as nominalization markers, similar to the nominalization markers *no* in Japanese and *de* in Chinese, which also head a relative clause (Chinese) and mark nominal modifiers in genitive and other noun modifier constructions (Japanese and Chi-

[2] See (20-22) above and Shibatani (2009) for further examples.

nese), and gloss it as NMZ, as we did in (23), in all its occurrence in the relevant constructions.

Because the identity of what are referred to by argument nominalizations is unspecified, the distribution of argument nominalizations is constrained in such a way that they typically occur in those contexts where the entity identification is provided or sought – in (a) relative clause constructions, where the head supplies the identification, (b) *wh*-questions that seek the identity of the entity referred to in the equation format (see (25a) below), (c) cleft-type constructions, in which the entity identification is made in the equation format (see (25b) below), (d) where the entity identification can be made from the context (see example (24) above), or (e) the entity referred to is generic or arbitrary (see the examples in (23) above).

(25) a. mən [dɛ jədəħ ʔəmʃɛn]
who NMZ.SG.M come-3SG.M.PER yesterday
'Who is the one who came yesterday?'

b. dɛ [jədəħ ʔəmʃɛn] ø ʕəli
NMZ.SG.M. come-3SG.M.PER yesterday COP Ali
'The one who came yesterday is Ali.'

In formal analysis, these constructions are often treated as involving extraction of an element from the clause – a wh-word in (25a) and the identifier noun phrase in (25b). There is really no basis for such analysis. These constructions are basically equational constructions of the type [A is B] involving two nominals, just like *Who is this?* and *This is Ali*. The only special feature with the expressions in (25) is that they carry the presupposition associated with the nominalization involved, namely that someone came yesterday, contrasting with the following where there is no such presupposition.

(26) a. mən jədəħ ʔəmʃɛn
who come-3SG.M.PER yesterday
'Who came yesterday?'

b. ʕəli jədəħ ʔəmʃɛn
Ali come-3SG.M.PER yesterday
'Ali came yesterday.'

6. From apposition to modification

Of the nominalizations discussed in this paper, perhaps the least obvious is the "genitive" nominalization, which creates a nominal out of a (pro)noun. But the wide use of a nominalization marker as a genitive marker, as in Chinese and Japanese, has been noted by Matisoff (1972). Essentially, what the "genitive" nominalization does is to create another noun referring to an entity "pertaining to" to what is referred to by a (pro)noun. Thus, Soqotri expression *dɛ ʕəli* means "an entity pertaining to Ali", where the precise interpretation of "pertaining to" is determined by context. The expression such as *ktɑːb dɛ ʕəli* is likely to have arisen from an appositive construction of the form [ktɑːb] + [dɛ ʕəli] with the meaning like 'the book, the thing pertaining to Ali'. When these two nominals are unified into a noun phrase unit as [[ktɑːb] [dɛ ʕəli]]$_{NP}$, the non-head element takes on a modification function resulting in the meaning such as 'the book pertaining to Ali'. Still the meaning of 'pertaining to' is indeterminate and the whole expression can mean the book possessed by Ali, the book written by Ali, the book about Ali, etc. The same applies to the modification constructions involving regular nouns such as [ktɑːb dɛ ħsɑːb] 'mathematics book' and [mɛssə dɛ χtɛːr] 'night rain'.

The development of modification function in a nominal element is seen in other [nominal + nominal] constructions as in noun compounds such as [[song][bird]] and [[night][train]] and phrasal forms such as [[bird] [singing in the bush]] and [[singing][bird]]. The development of the modification function here is perhaps due to the constraint that a single noun phrase refers to only a single entity. We would analyze Soqotri relative clauses and adjectival modification constructions the same way, namely modification by the nominalizations below is the function of the NP constructions into which they enter:

(27) a. [dɛʃ [təʕɑːɣɛn birhɛ]]$_{NP}$
NMZ.SG.F 3SG.F-love children
'(feminine) one who loves children'

b. [ʁɑːjɛh [dɛʃ [təʕɑːɣɛn birhɛ]]]$_{NP}$
woman NMZ.SG.F 3SG.F-love children
'the woman who loves children'

(28) a. [dɛ [ħɑːr]]$_{NP}$
NMZ.3SG.M black.3SG.M
'one that is black'
b. [ʁəʃ [dɛ [ħɑːr]]]$_{NP}$
man NMZ.SG.M black-SG.M
'the black man'

7. Lexical nominalization and grammatical nominalization

As mentioned above, Soqotri shows some measure of productivity in lexical argument nominalization, which derives nouns from perfective verb forms referring to different types of semantic roles.

(29) Lexical nominalization
a. Agent/subject nominalization
[ʃɛʕrɛk]$_V$ → [**m**ʃɛʕrɛk]$_N$
'fished' 'fisherman'
[s'ɛːli]$_V$ → [**m**s'əlli]$_N$
'prayed' 'prayer/one who prays'
[ʃɛk'ɑːbit]$_V$ → [**mi**ʃɛk'ɔːbɒt]$_N$
'learned' 'learner'
[nk'ɛːnəʕ]$_V$ → [**m**ənkɛːnəʕ]$_N$
'went mad' 'madman'

b. Patient/object lexical nominalization
[lɑːtəʁ]$_V$ → [liːtəʁ]$_N$
'killed' 'one who was killed'
[ħɔːbɛs]$_V$ → [ħiːbɛs]$_N$
'jailed' 'one who was jailed/jail bird'
[ʔɛːɟɛ] → [ʔiːɟɛ]$_N$
'beat' 'one who was beaten'

c. Instrumental nominalization
[ʃəħlɛf]$_V$ → [**m**ʃɛħlɛf]$_N$
'sharpened' 'sharpener'
[k'ɒdɛf]$_V$ → [**m**ɒk'dɛf]$_N$
'rowed' 'oar'
[ʁɒrɛf]$_V$ → [**m**ɒʁrɛf] $_N$
'took water 'a big glass'
'with a glass'

While lexical nominalization of the above type involves change in vocalic melody morphemes, grammatical nominalization does not – it simply marks what is nominalized by the nominalization marker *dɛ*. Grammatical nominalization also yields argument nominalization as lexical nominalizations. Formal treatment of this process has not been studied, but we may characterize the process in the following way:

(30) Grammatical nominalization
a. Agent/subject nominalization
[ʁəɟ tɔːɟɛr ʔɒz] → dɛ [Ø$_i$ tɔːɟɛr ʔɒz]$_i$
man buy.3SG.M.PER goat NMZ buy.3SG.M.PER goat
'The man bought a goat.' 'the one who bought a goat'
[ʁəɟ ħɑːr] → dɛ [Ø$_i$ ħɑːr]$_i$
man black-SG.M NMZ black.SG.M
'The man is black.' 'a black one'

b. Patient/object nominalization
[ʁəɟ tɔːɟɛr ʔɒz] → dɛ [ʁəɟ tɔːɟɛr Ø$_i$]$_i$
man buy.3SG.M.PER goat NMZ man buy.3SG.M.PER
'The man bought a goat.' 'the one that the man bought'

c. "Genitive" nominalization
[sɛh] → dɛ [sɛh]$_i$ [χtɛːr] → dɛ [χtɛːr]$_i$
she NMZ she night NMZ night
'she' 'hers' 'night' 'of the night'

The essential point is that grammatical nominalization produces a nominal, which is coindexed with the argument nominalized (argument nominalization) or which simply takes on a new referential status ("genitive"

nominalization), and which, together with the nominalization marker *dɛ*, forms a noun phrase or D(eterminer) P(hrase), referring to an entity that is characterized by the role of the derived nominal as an agentive or patientive event participant, or an entity pertaining to what is denoted by the noun being nominalized. In other words, these expressions do not involve a deleted lexical head or an empty pronominal head as in the widely practiced conventional analysis (see above).

As in other languages, some grammatical nominalizations have been lexicalized. For example, *dɛ laːtəʁ* can mean 'a killer' in addition to the grammatical nominalization reading of 'a one who killed'. The unmarked patient/object nominalization forms in (29b) can be preceded by the nominalization marker *dɛ* and become indistinguishable from grammatical nominalizations of the same form. It thus appears that the distinction between lexical and grammatical nominalization is blurred in certain expressions, indicating that a fair number of *dɛ* grammatical nominalizations have become lexicalized. If this conjecture on the current state of the *dɛ* nominalization in Soqotri is correct, we can further surmise that lexical agentive/instrumental nominalization marked by *m-* or the unmarked patientive nominalizations may have been once productive grammatical nominalizations, as the parallel involvement of *m-* in agentive/instrumental nominalization and genitive nominalization indicates (see the earlier remark on this). Thus, the current grammatical nominalizations marked by *dɛ* – at least those that refer to what are deemed nameworthy entities – may become lexicalized eventually, when a new grammatical nominalization is likely to be invented.

Soqotri possesses several other types of nominalization used as verb complements and other types of subordinate expressions, including some types of participial expressions. These and other possible types of nominalization require a separate treatment from the analysis advanced above. But the essential point of this paper, namely that (grammatical) nominalization plays a very important role in grammar extends beyond the scope of this paper.

8. Conclusion

Both descriptive and theoretical linguists have tended to be preoccupied with form and have often failed to recognize some salient phenomena simply because they are not accompanied by some formal marking. In

the case of nominalization, grammarians' preconception centers on derivational morphology. Even when clausal nominalization is talked about, some non-finite features are sought after such as participial or gerundive forms of verbs and/or genitive marking of the subject nominal. Indeed, the typical initial reaction of most practicing linguists is negative when asked whether or not the complement clause in *John thinks that Bill is honest* is nominalized. Nominalization, much like other notions such as relativization and causation, is a functional, not a morphological or formal, notion, referring to the creation of referring expressions. Grammatical nominalization we discussed in this paper creates nominal expressions referring to entities characterized in terms of an event in which they are involved as an agent/subject, patient/object, etc. Since the entities referred to are characterized in terms of events, the process of grammatical nominalization is perfectly compatible with tense/aspect marking and other finiteness features, since the speaker may want to characterize the relevant entity in terms of its having bought a book yesterday, for example, e.g., as in *(the one) who bought a book yesterday*.

Ironically, the type of grammatical nominalization we examined in this paper involves a clear formal marking, namely the particle *dɛ* and its variants that occur in front of the nominalized form. The problem here and elsewhere (see Shibatani 2009) was the wrong perspective both typologists and theoretical linguists have had about grammatical nominalizations; namely, looking at them from the point of view of relativization, as if the relevant nominalizations were derivatives of relative clauses. In this paper, we have endeavored to clarify these two problems inherent in the past studies of nominalization and relativization, arguing for the nominalization-based analysis of relativization and other types of noun modification constructions. A proper understanding of the nature of nominalization also explains why nominalizations are used in relativization and occur in other environments such as cleft-constructions and cleft-based wh-questions.

References

Adams, Douglas Q. (1972). Ancient Greek. In Paul M. Peranteau, Judith N. Levi & Gloria C. Phares (eds.), *The Chicago Which Hunt: Papers from the Relative Clause Festival*, 9-22. Chicago: Chicago Linguistic Society.

Andrews, Avery (2007). Relative clauses. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*. Volume 2, (Second edition), 206-236. Cambridge: Cambridge University Press.

Comrie, Bernard & Sandra A. Thompson (2007). Lexical Nominalization. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, 334-381. Volume 3, (Second edition). Cambridge: Cambridge University Press.

Deutscher, Guy (2009). Nominalization and the origin of subordination. In T. Givón & Masayoshi Shibatani (eds.), *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution*, 199-214. Amsterdam: John Benjamins.

Dryer, Matthew S. (2005). Order of relative clause and noun. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 366-367. Oxford & New York: Oxford University Press.

Dryer, Matthew. S. (2007). Noun phrases. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, 151-205. Volume 2, (Second edition). Cambridge: Cambridge University Press.

Huang, Chenglong (2008). Relativization in Qiang. *Language and Linguistics* 9.4: 735-768.

Keenan, Edward L. (1985). Relative clauses. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, 141-170. Cambridge: Cambridge University Press.

Lehmann, Christian (1986). On the typology of relative clauses. *Linguistics* 24: 663-680.

Makhashen, Khaled Awadh (2009). *Lexical and Syntactic Aspects of the Soqotri Dialect of Galansiyah*. PhD dissertation, Universiti Sains, Malaysia.

Makhashen, Khaled Awadh, Munir Shuib & Salasiah Che Lah (2008). The noun phrase structure in Soqotri. Paper presented at the *International Conference on Moribund Languages andCultures* 14-16 October 2008, University of Malaya.

Matisoff, James A. (1972). Lahu nominalization, relativization, and genitivization. In John P. Kimball (ed.), *Syntax and Semantics*, 237-257. Volume 1. New York: Academic Press.

Matisoff, James A. (1973). *The Grammar of Lahu*. Berkeley: University of California Press.

Naumkin, V.V. (1998). Personal pronouns in Socotran folklore texts. In *the proceedings of the seminar for Arabian studies* 28, 229-232.

Noonan, Michael (1997). Versatile nominalizations. In Joan Bybee, John Haiman & Sandra Thompson (eds.), *Essays on Language Function and Language Type. In Honor of T. Givón*. Amsterdam/Philadelphia: John Benjamins.

Shibatani, Masayoshi (2009). Elements of complex structures – where recursion isn't: The case of relativization. In T. Givón & Masayoshi Shibatani (eds.), *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution*, 163-198. Amsterdam: John Benjamins.

Simeone-Senelle, Marie-Claude (1997). The Modern South Arabian languages. In Robert Hetzron, (ed.) *The Semitic languages*, 378-423. London: Routledge.

Sneddon, James N. (1996). *Indonesian: A Comprehensive Grammar*. London: Routledge.

Weber, John D. (1983). *Relativization and Nominalized Clauses in Huallaga (Huanuco) Quechua*. Berkeley: University of California Press.

Wrona, Janick (2008). The Nominal and Adnominal forms in Old Japanese: Consequences for a reconstruction of pre-Old Japanese syntax. In Bjarke Frellesvig & John B. White (eds.), *Proto-Japanese*, 199-221. Amsterdam:John Benjamins.