Noun or quantifier?
Examining the semi-lexicality of Polish and English quantificational expressions

Heidi Klockmann (hklockmann@gmail.com)
Utrecht University

CRISSP seminar, Brussels
April 19, 2018

1. Introduction

Languages use many expressions to indicate quantity in the nominal domain:

- Cardinal numerals: one, two, three, four, Polish jeden ‘one’, dwa ‘two’, trzy ‘three’
- Relative quantifiers (context dependent ranges): many, few, little, much
- Absolute quantifiers (fixed ranges/values): several, a couple, a dozen
- Noun-like quantifiers: a bunch, a ton, a lot, a dozen, a couple, Dutch een hoop ‘a lot’

There have been many views on how to treat the category of quantificational expressions:

- (Some) numerals are adjectives (Corbett 1978; Zabbal 2005; Stavrou & Terzi 2008), which some take to modify a silent noun (Zweig 2005)
- (Some) numerals are nouns (Jackendoff 1977; Corbett 1978; Zabbal 2005; Corver & Zwarts 2006; Ionin & Matushansky 2006; Stavrou & Terzi 2008)
- Numerals can be any lexical category: noun, verb, or adjective (Donohue 2005)
- (Some) quantifiers are adjectives (Solt 2009), which some take to modify a silent noun (Kayne 2005)
- Quantifying nouns in pseudopartitives are nouns (Jackendoff 1977; Selkirk 1977; Abney 1987).
- Quantifying nouns in pseudopartitives are semi-lexical heads (van Riemsdijk 1998), part of a measure phrase or classifier phrase (Stavrou 2003; Alexiadou, Haegeman, and Stavrou 2008; Stickney 2009).

A common observation, inherent in many of the analyses mentioned above, is that quantificational expressions often resemble lexical categories, and might even be instantiations of lexical categories.

This talk is concerned with the category of quantificational expressions, focusing on two case studies: Polish numerals and English quantificational nouns (bunch, lot, ton).

- What category are Polish numerals and English quantificational nouns?
- Why do they show a resemblance to lexical nouns and adjectives?
- What does this imply for the syntactic representation of quantifying words?
Roadmap:
- Quantity in the nominal domain: What is the structure of a DP and where do quantificational expressions fit in?
- Semi-lexicality: The phenomena studied here appear to be semi-lexical – what is semi-lexicality and what does it add to our understanding?
- Case study #1: Polish numerals, with a focus on numeral 1000
- Case study #2: English quantificational nouns (Q-nouns), *bunch, ton, lot*

2. On the notion of quantity in the nominal domain

2.1 Quantity as a universal interpretive function

Cinque (2005) and Abels and Neeleman (2006) show that cross-linguistically, there seems to be a single order of Merge in the nominal domain (with the existing set of deviations derivable via movement), namely:

(1) Demonstrative > Numeral > Adjective > Noun

This suggests a **functional sequence** in the nominal domain, which includes **quantity**.

Wiltschko (2014) defends the **Universal Spine Hypothesis**: Languages share a universal spine (each piece having a dedicated interpretive function), but can differ in their implementation (i.e. the categories they use). Her spine is represented below:

(2) Linking > Anchoring > Point-of-View > Classification
    
    CP/KP    IP/DP    AspP/φP    vP/nP

    **CLASSIFICATION**: Introduces and classifies some root as an event (verb) or individual (noun)
    **POINT-OF-VIEW**: Adds a viewpoint with respect to the event or individual
    **ANCHORING**: Anchors the event or individual to the utterance
    **LINKING**: Links the event or individual to the existing (discourse) structure

Hachem (2015) builds on this, but proposes “universal distinction domains” (a set of projections fulfilling each interpretive function). She also adjusts the content of the spine:

(3) The hierarchy of Universal Distinction Domains (Hachem 2015: 57)
    
    Linking > Anchoring > Quantity > Classification > Identification
    
    PP/KP    DP    QP    φP    √

    **IDENTIFICATION**: Converts a chunk of encyclopedic knowledge into a root for use in the syntactic structure
    **CLASSIFICATION**: Classifies the root as an individual (via φP, e.g. number and gender)
QUANTITY: Determines the size of quantities (via QP)
ANCHORING: Anchors the individual to the utterance (via DP)
LINKING: Links the individual to the existing (discourse) structure (via PP / K[ase]P[hrase])

Implications of this style of approach:

- Quantity/quantification is universally available (forming a part of the spine).
- Languages can use language-specific categories, and therefore, we do not expect a universal category for implementing quantity.
- These language-specific categories may involve multiple projections (i.e. the universal distinction domain) which accomplish the function.

- This predicts that we should find plenty of variation in the types of expressions used to express quantity (= category), and potentially, in the types of quantities expressed (e.g. the set of numerals, ranges of quantifiers, etc.).

General hypothesis (which this talk attempts to validate):

- Languages can innovate quantificational expressions, according to the needs of the linguistic community.
- When creating a new quantificational expression, languages will make use of existing resources, e.g. lexical categories, or pre-existing quantificational expressions. In the first case, this leads to a resemblance to lexical categories.
- These can eventually grammaticalize into a language-specific category, expressing the quantity function.

2.2 DP structure and quantity

Drawing on the Universal Spine Hypothesis (Wiltschko 2014, adapted in Hachem 2015), and exoskeletal approaches to DP structure (Borer 2005, de Belder 2011), I assume the following set of skeletal projections and structures:

- **DP**: Definiteness phrase (Abney 1987), or determiner phrase (Lyons 1999)
- **QP**: Quantifier phrase (e.g. Löbel 1989; Giusti 1991; a.o.)
- **#P**: Number phrase (Ritter 1991, 1992; a.o.)
- **γP**: Gender phrase (Picallo 1991; superficially similar to Lowenstamm 2008)
- **Root**: the lexical root

Each of these projections are part of a particular domain, and hence, could theoretically be decomposed into more projections (e.g. a high and low DP as in Zamparelli 2000).

Note that not all projections need be present in all structures, e.g. if no quantity is expressed, QP may be absent (and likewise, with a mass noun, #P may be absent).
4. Skeletal structure of an English DP:

\[
\begin{align*}
\text{DP} &= \text{ANCHORING Domain} \\
D &\quad \text{QP} = \text{QUANTITY Domain} \\
Q &\quad \#P = \text{CLASSIFICATION Domain} \\
\# &\quad \sqrt{\gamma} = \text{IDENTIFICATION Domain (root)}
\end{align*}
\]

5. Skeletal structure of a Polish DP:

\[
\begin{align*}
\text{DP} &= \text{ANCHORING Domain} \\
D &\quad \text{QP} = \text{QUANTITY Domain} \\
Q &\quad \#P = \text{CLASSIFICATION Domain} \\
\# &\quad \gamma P \sqrt{\gamma} = \text{IDENTIFICATION Domain (root)}
\end{align*}
\]

The realization of the quantity domain (or rather QP) will be the main focus of this talk.

How does a particular lexical item “know” that it belongs in the quantity domain?

- I will adopt the convention of assigning quantificational expressions a [Q] feature.
- This, in effect, avoids the question, but it will suffice for now, until a more explanatory answer can be found.

3. Numerals and quantificational nouns (Q-nouns) as semi-lexical

Are numerals and Q-nouns lexical or functional?

- They are **not** fully functional.

Numerals are an open class. We can make new numerals easily:

6. Fictitious quantities: jillion, bajillion, zillion, gazillion

7. Very high quantities: googol \((10^{100})\), googolplex \((10^{10^{100}} \text{ ‘ten to the ten to the hundred’})\), novemdecillion \((10^{60})\), vigintillion \((10^{63})\) (see wikipedia for more)

8. Developments in the (high) base numerals hundred, million, billion, and trillion (based on the Corpus of Historical American English (Davies 2010–))
   a. Hundred: Rises from 224 hits in 1810 (189.64 hits per million words) to 5457 hits in 1860 (319.97 hits per million words), then declines
   b. Million: Rises from 24 hits in the 1810s (20.32 hits per million words) to 9496 hits in the 2000s (321.16 hits per million words)
c. *Billion*: No hits before 1840; takes off in the 1900s

d. *Trillion*: No hits before 1840; takes off in the 1900s (especially the 1980s)

Numerals *seem* to be a closed class, but this follows from the fact that we have already named the most useful quantities. If needed, new numerals can be added to the system.

Given that we can create new numerals, and that numerals can theoretically express an infinity of values, it seems undesirable to describe them with purely functional features.

(Some) numerals seem to be derived from nouns:

- Numeral 5 in Polish can be traced back to a noun meaning ‘hand,’ and numeral 10 to ‘the whole of hands’ (Jagodziński 2016), presumably based on the analogy that a hand has five fingers (and two hands makes ten fingers).
- Do numerals have roots? Kernels of meaning expressing ‘oneness’, ‘twoness’, ‘threeness’, etc.? This would argue against their functionality.

Q-nouns are also an open class. Many nouns can be used to express quantity in a construction similar to what is found with *lot*, *ton*, and *bunch*. These examples were extracted from the Corpus of Contemporary American English (COCA) (Davies 2008-):

(9) A flood of memories
(10) A parade of witnesses
(11) A torrent of words
(12) A sea of faces
(13) A flurry of lawsuits
(14) A mass of contradictions
(15) A chorus of boos
(16) A mob of reporters
(17) A multiplicity of voices
(18) A barrage of insults
(19) A cascade of problems

- Do Q-nouns have roots? These examples appear to involve the lexical semantics of the quantifying noun, suggesting the presence of a root.

Q-nouns also seem to be derived from nouns.
- *Ton*: quantificational meaning ‘many’; lexical meaning ‘2000 pounds’
- *Bunch*: quantificational meaning ‘many’; lexical meaning ‘a number of things, bundled together in some way’

They are **not** fully lexical.

Numerals and Q-nouns show properties which differ from lexical categories. For example (with plenty more to come soon), agreement cannot target a Q-noun, instead targeting the quantified noun:
(20) Verbal agreement, SV order:
   a. A lot of people were/*was invited to the party.
   b. A ton of people were/*was invited to the party.
   c. A bunch of people were/*was invited to the party.
   d. Lots of power is/*are necessary to cool the building.
   e. Tons of power is/*are necessary to cool the building.

(21) Inverted yes-no questions, VS order:
   a. Were/*was a lot of people invited to the party?
   b. Were/*was a ton of people invited to the party?
   c. Were/*was a bunch of people invited to the party?
   d. Is/*are lots of power necessary to cool the building?
   e. Is/*are tons of power necessary to cool the building?

- They are “semi-lexical.”

   “Certain lexical items display ambiguous behavior: they share properties with lexical categories and at the same time they display functional characteristics.”

   (Corver and van Riemsdijk 2001: 3)

**Approach to semi-lexicality** (Klockmann 2017)

What does it mean to be **lexical**?

- Lexical categories include nouns, verbs, and adjectives.
- **Roots** have been identified at the core of nouns, verbs, and adjectives (Halle and Marantz 1993; Marantz 1997; Borer 2005).
- Under exoskeletal approaches, where a root is inserted determines its category. A root becomes a noun when inserted in the structure associated with a noun.
- Therefore: lexical implies the presence of a root.

What does it mean to be **functional**?

- Syntactic features refer to grammatical notions like tense, number, animacy, the EPP, agreement probes, etc.
- Syntactic features drive syntactic computation (case, agreement, movement, Merge, etc.); they are presumably the core of functional categories.
- Therefore: functional implies the presence of a syntactic feature.

What does it mean to be **semi-lexical**?

- Semi-lexicality is often cited as the combination of lexical and functional in a single lexical item. Taking this very literally...
- **Hypothesis**: Semi-lexicality is what occurs when a root is lexically specified for a syntactic feature.
4. Case study #1: Polish numerals

Polish cardinal numerals can be divided into four classes, according to their morphosyntactic properties, and the categories they resemble:

- Numeral 1 (adjective)
- Numerals 2, 3, 4 (adjective-like and noun-like)
- Numerals 5-10, 100 (adjective-like and noun-like)
- Numeral 1000 (noun-like)

4.1 Numeral 1000

4.1.1 Data: Comparing 1000 to a lexical noun

Numeral 1000 resembles a lexical noun in its morphological paradigm (compared here to a similar-sounding noun miesiąc ‘month’):

\[(22)\] Table 1: Paradigm of 1000 and lexical noun miesiąc ‘month’

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PYL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>tysiąc</td>
<td>tysiąc-e</td>
</tr>
<tr>
<td></td>
<td>miesiąc</td>
<td>miesiąc-e</td>
</tr>
<tr>
<td>ACC</td>
<td>tysiąc-a</td>
<td>tysiąc-y</td>
</tr>
<tr>
<td></td>
<td>miesiąc-a</td>
<td>miesiąc-y</td>
</tr>
<tr>
<td>GEN</td>
<td>tysiąc-owi</td>
<td>miesiąc-om</td>
</tr>
<tr>
<td></td>
<td>miesiąc-owi</td>
<td>miesiąc-om</td>
</tr>
<tr>
<td>DAT</td>
<td>tysiąc-u</td>
<td>miesiąc-ach</td>
</tr>
<tr>
<td></td>
<td>miesiąc-u</td>
<td>miesiąc-ach</td>
</tr>
<tr>
<td>LOC</td>
<td>tysiąc-em</td>
<td>miesiąc-ami</td>
</tr>
<tr>
<td>INST</td>
<td>tysiąc-em</td>
<td>miesiąc-ami</td>
</tr>
</tbody>
</table>

It also resembles a lexical noun in requiring genitive on the quantified noun in all case contexts (compared here to klucz ‘flock (lit. key)’):

\[(23)\] Tysiąc ptaków / Klucz ptaków
Thousand.NOM birds.Gen / Key.NOM birds.Gen
‘A thousand birds / A flock of birds (flying in a V)’

\[(24)\] Z \{ tysiącem ptaków / kluczem ptaków \}
With[INST] \{ thousand.INST birds.Gen / key.INST birds.Gen \}
‘With a thousand birds / a flock of birds (flying in a V)’

Where it differs concerns agreement patterns. With a lexical noun, the verb (in the past tense) agrees for person, number, and gender.
(25) Chłopcy spali i dziewczyny spały.
Boys.M.PL.NOM slept.3.v.PL and girls.F.PL.NOM slept.3.nv.PL
‘The boys slept, and the girls slept.’

I report on numeral 1000. The results below are gathered from the Polish National Corpus (using the balanced NKJP subcorpus of 300m words) (Przepiórkowski et al. 2011), and were tested against the grammaticality judgments of five native speakers.

- When numeral 1000 is bare and singular, verbal agreement is default (n=5):¹

(26) Przez ciebie tysiąc kijów spadło na mój grzbiet!
By you thousand sticks.GEN fell.N.SG on my back
‘Because of you, a thousand sticks fell on my back!’

(27) Wokół niego tysiąc gwiazd migotało na niebie …
Around him thousand stars.GEN flickered.N.SG on heavens
‘Around him, a thousand stars flickered in the heavens…’

- When 1000 is bare and plural, verbal agreement is default or successful; if successful, it targets the numeral (n=1 of 1):²

(28) Tysiące pytań przechodziło/przechodziły mi do głowy.
Thousands questions.GEN came.N.SG /came.NV.PL me to head
‘Thousands of questions came to my head.’

- If numeral 1000 is preceded by an agreeing pre-modifier (adjective, demonstrative), verbal agreement is successful (targets the numeral), and if it is preceded by a non-agreeing pre-modifier, verbal agreement is default (n=5):³

Agreeing pre-modifier and verb:

(29) …i dobry tysiąc kamieni przemknął nam
and good.M.SG thousand.NOM rocks.GEN flew.M.SG us
nad karkami.
over necks
‘and a good thousand rocks flew over our heads (lit. necks).’

¹ The following searches were conducted in the 300m word corpus.
   i. Default agreement: [pos!=adj] tysiąc [pos=verb & number=sg & gender=m3]
   ii. Agreement: [pos!=adj] tysiąc [pos=verb & number=pl]
There were 88 default agreement hits (most unverified), and 7 agreement hits (verified not to be genuine).

² The following searches were conducted in the 300m word corpus:
   i. Agreeing verbs: [base=tysiąc & number = pl & case=nom] [pos=verb & number=pl] (103 hits, unverified)
   ii. Non-agreeing verbs: [base=tysiąc & number = pl & case=nom] [pos=verb & number=sg] (190 hits, unverified)

³ The following searches were conducted in the 300m word corpus:
   i. Agreeing premodifiers: [pos=adj & gender=m3 & number=sg] tysiąc (187 hits, unverified)
   ii. Non-agreeing premodifiers: [pos=adj & gender=m3 & number=pl] tysiąc (33 hits, unverified)
(30) Do Albanii dotarł pierwszy tysiąc żołnierzy. 
To Albania reached.first.M.SG thousand.NOM soldiers.GEN 
sił międzynarodowych. 
force.GEN international.GEN
‘The first thousand soldiers from the international forces reached Albania.’

(31) Ponieważ ten tysiąc zawierał prawie całą śmieiankę polskiego towarzystwa...
Because DEM.M.SG thousand.NOM contained.M.SG almost all.ACC 
śmietankę polskiego towarzystwa... 
cream.ACC Polish.GEN society.GEN
‘Because this thousand contained almost all the cream of Polish society’

Note: The set of pre-modifiers which can agree with the numeral is subject to speaker variation; more research is needed on this.

Non-agreeing pre-modifier and verb:

(32) Te tysiąc złotych przeznaczone było na jakiś cel i już zostało wydane. 
DEM.NV.PL thousand.NOM gold.GEN designed.NV.PL was.N.SG for some purpose and already was.N.SG spent.NV.PL
‘These thousand gold(Polish currency) were designed for some purpose and have already been spent.’

• Verbal agreement cannot be default if the pre-modifier agrees (n=4, of 5): ^4

(33) %Dodatkowy tysiąc osób dał-o-by się upchnąć w prywatnych pensjonatach. 
Additional.M.SG thousand people.PL allow-N.SG(DEF)-COND SIE push.INF in private.LOC pensions.LOC
‘An additional one thousand people would allow themselves to be pushed into private pensions.’

• Verbal agreement cannot be successful if the pre-modifier does not agree (n=5): ^5

(34) *Dodatkowe tysiąc osób dał-by się upchnąć w prywatnych pensjonatach. 
Additional.NV.PL(DEF) thousand people.PL allow.M.SG-COND SIE push.INF in private.LOC pensions.LOC

• Pre-modifiers can agree with the quantified noun, with default verb agreement (n=5):

---
^4 The search was expanded to the 1800m word corpus, given that no relevant examples surfaced in the 300m word corpus. The search was the following:
  i. [pos=adj & gender=m3 & number=sg] tysiąc [pos=subst] [pos=praet] (1 genuine hit)

^5 The search was also expanded to the 1800m word corpus. The search was the following:
  i. [pos=adj & number=pl] tysiąc [pos=subst] [pos=praet] (0 genuine hits)
Kolejnych tysiąc osób przyszło osobiście albo zadzwoniło.

‘Another thousand people came in person or called.’

Ze tych tysiąc górników od razu poszło się...

‘…that a thousand miners went at once (to…)’

There are two patterns: either everything, pre-modifiers and verbs, agrees with the numeral for masculine singular (Pattern 1), or nothing agrees with the numeral, instead targeting the quantified noun or surfacing with default morphology (Pattern 2).

4.1.2 Analysis: Thesemi-lexicality of numeral 1000

Puzzle 1: Why is a pre-modifier a condition for agreement?

Idea (not worked out): There needs to be enough structure to host the pre-modifier; the inclusion of this structure allows the numeral the space to project phi-features, this making it an agreement target.

Puzzle 2: Why must the verb agree if the pre-modifier has agreed?

Solution: Successful agreement marks the presence of phi-features on the numeral (masculine singular or plural). If the numeral has phi-features, it is a target for agreement, and thus, will control the adjective and verb.

Default agreement marks phi-feature absence on the numeral. The numeral is phi-feature deficient and the noun genitive; there is no target for successful verb agreement, though the pre-modifier can target the noun (genitive) or numeral (default).
Conclusion: The numeral has two realizations: one with phi-features and one without.

A: Numeral 1000 is a lexical noun with γ and #. (= successful agreement)
B: Numeral 1000 is a deficient noun, with only # (no γ). (=default agreement)

We can model this in the lexical entry of the numeral, assuming a root meaning something like ‘thousand-ness’, which is specified for phi-features.

Lexical 1000 (similar to other lexical nouns which are specified for gender):

(41)  
tysiąc ‘1000’: [√1000, m[i], Q]  specified as masculine inanimate

Deficient 1000 (with a “negative” feature to mark the obligatory absence of gender).

(42)  
tysiąc ‘1000’: [√1000, ¬γ, Q]  specified for the absence of gender

Since aside from phi-feature specifications, there seem to be no other reasons to treat these separately, we can combine them, with an optional gender feature:

(43)  
tysiąc ‘1000’: [√1000, (¬)m[i], γ, Q]

Further evidence for the claim that numeral 1000 can surface in two syntactically different forms comes from (a) distributive marker po and (b) coordination.

(a) Distributive marker po requires locative case on lexical nouns (see Przepiórkowski 2010, 2013, Przepiórkowski and Patujek 2013):

(44)  
Dalam im po jabłku / *jabłko
Gave.1SG.F them DIST apple.LOC / *apple.ACC
‘I gave them an apple each.’

“Numerals” take the case of the case context with po:

(45)  
Dalam im po pięć jabłek / *pięciu jabłkach.
Gave.1SG.F them.DAT PO five.ACC apples.GEN / five.LOC apples.LOC
‘I gave them five apples each.’

Numeral 1000 shows optionality – it either surfaces as locative like a lexical noun (= has γ), or in the case of the case context like a numeral (= does not have γ):

(46)  
Dalam im po tysiąc / tysiącu jablek.
Gave.1SG.F them DIST thousand.ACC / thousand.LOC apples.GEN
‘I gave them a thousand apples each.’
(b) Conjoined verbs should not show two agreement patterns with one numeral 1000:

(47) a. *Całe tysiąc ptaków jadł i spało.
     Whole.DEF thousand birds.GEN ate.M.SG and slept.N.SG
     ‘A whole thousand birds ate and slept.’

b. *Całe tysiąc ptaków spało i jadł.

(48) a. *Cały tysiąc ptaków jadł i spało.
     Whole.M.SG,NOM thousand birds.GEN ate.M.SG and slept.N.SG
     ‘A whole thousand birds ate and slept.’

b. *Cały tysiąc ptaków spało i jadł.

Structure: Numeral 1000 appears to have its own classification domain (given the fact that it has phi-features), within the quantity domain of the quantified noun:

(49) \[\begin{array}{c}
\text{DP = ANCHORING domains I \& II} \\
\text{D} \\
\text{\#P = QUANTITY domain I / CLASSIFICATION domain II} \\
\text{\#_{SG,PL}} \\
\text{(γP)} \\
\text{(γM[I])} \\
\text{√P} \\
\text{√1000Q} \\
\text{\#P = CLASSIFICATION domain I} \\
\text{\#_{PL}} \\
\text{γP} \\
\text{γ} \\
\text{√}
\end{array}\]

Numeral 1000 can also combine with other numerals and quantifiers, suggesting it can host a quantity domain of its own (again inside the quantity domain of the noun):

(50) Dwa tysiące ptaków
     Two.M.NOM thousands.NOM birds.GEN
     ‘Two thousand birds’

(51) Pięć tysięcy ptaków
     Five.NOM thousands.GEN birds.GEN
     ‘Five thousand birds’

(52) Kilka / wiele / ile tysięcy ptaków
     Several / many / how many thousands.GEN birds.GEN
     ‘Several thousand birds / (how) many thousands of birds’

If this is more or less the structure of a noun, then it predicts that adjectives can surface between the numeral and what quantifies it. This is true (corpus examples verified):
Tych kilkanaście nędznych tysięcy dolarów ustawiało ich finansowo …

‘Those miserable several/many thousands of dollars ($11,000-$19,000) set them up financially.’

Kilka kolejnych tysięcy świętowało udział ich faworytów w finale na ulicach hiszpańskiego miasta.

‘The next several thousand celebrated the participation of their favorites in the final on the streets of a Spanish city.’

Gdyby panu, nie daj Bóg, coś się stało,

‘If God forbid anything happened to your additional thirty thousand’

Structurally, this implies a classification and quantity domain for the numeral within the quantity domain of the noun, as depicted below:

This accounts for the noun-like nature of the numeral: it has the structure of a lexical noun (=noun-like), but occurs in the quantity domain of another noun (= numeral).

Modeling the agreement facts: I've adopted a feature sharing (Frampton and Gutmann 2000; Pesetsky and Torrego 2007, a.o.) and dependent case approach (Marantz 1991; Baker 2015; Levin 2015, a.o.) (see Klockmann 2017 for discussion).
**Lexical 1000:**
- Numeral and noun both associated with phi-complete phi-bundles ($\varphi_1$ and $\varphi_2$).
- $\varphi_1$ (numeral) functions as an intervener, blocking agreement with $\varphi_2$ (noun).
- Agreement by all probes is with $\varphi_1$ (numeral).

(57)

$\text{TP} \rightarrow \text{DP} \rightarrow \text{T'}$

\[
\begin{array}{c}
\text{D} \quad \#P \\
\text{[π, } _γ ᾱ _#]\quad \gamma P \\
\text{s}_γ / \text{p}_γ \\
\varphi_1 \\
\text{[NOM]} \\
\sqrt{1000Q} \quad \#P \\
\varphi_2 \\
\text{[GEN]} \\
\end{array}
\]

**Deficient 1000:**
- Numeral and noun both associated with phi-bundles ($\varphi_1$ and $\varphi_2$).
- $\varphi_1$ (numeral) is phi-incomplete, and $\varphi_2$ (noun) is phi-complete, but genitive.
- **Assumption:** Agreement probe can “use” $γ$ or # to probe (idea from Ionin and Matushansky 2016), with the result that either $\varphi_1$ is the target (probing on #) or $\varphi_2$ is the target (probing on $γ$).
- Verb will be default (needs something phi-complete and uncased); pre-modifier will be default ($\varphi_1$) or match the # and $γ$ of its target ($\varphi_2$).

(58)

$\text{DP} \rightarrow \text{D} \rightarrow \#P$

\[
\begin{array}{c}
\text{D} \quad \#P \\
\text{[γ, } _γ ᾱ _#]\quad \gamma P \\
\text{s}_γ / \text{p}_γ \\
\varphi_1 \\
\sqrt{1000Q} \quad \#P \\
\varphi_2 \\
\text{[GEN]} \\
\end{array}
\]
4.1.3 Numeral 1000 in transition

A property of numeral 5 is that, like a lexical noun, the quantified noun must appear in the genitive case:

(59) Pięć mądrych dziewczyn
    Five.NV smart.GEN girls.GEN
    ‘Five smart girls

Unlike a lexical noun, agreement is always default:

(60) Pięć ptaków spało.
    Five.NOM birds.GEN slept.N.PL
    ‘Five birds slept.’
(61) Ptaki spały.
    Birds.M.PL.NOM slept.N.PL
    ‘Birds slept.’

And, the genitive on the quantified noun disappears in oblique case contexts:

(62) Z pięcioma studentami
    with[INST] five.INST students.INST
    ‘with five students’
(63) Ze studentem fizyki /*fizyką
    with[INST] student.INST physics.GEN /*physics.INST
    ‘with a student of physics’

Historically, though, numeral 5 behaved very much like lexical 1000 today. It could control agreement:

(64) Pięć lat minęła.
    Five.F.SG years.GEN passed.F.SG
    ‘Five years passed.’
(65) Jako minęła dziesięć lat
    As passed.F.SG ten.F.SG years.GEN
    ‘As ten years passed.’ (Dziubała-Szrejbrsowska 2014: 103, ex. 132)

And, even in oblique case contexts, the quantified noun was genitive:

(66) Z piąci-ą-dziesiąt synów
    With[INST] five-GEN-ten.GEN.PL sons.GEN
    ‘With fifty sons’ (Miechowicz-Mathiasen 2012: 7, ex. 1b)

Observation: Numeral 1000 is becoming more like numeral 5 (p.c. Ewa Willim):

- It sometimes shows agreement (= lexical noun), but sometimes default agreement (= numeral).
While most speakers need a genitive on the quantified noun in all case contexts (=
lexical noun), some allow oblique case on the quantified noun (= numeral).

(67) I rozmawiamy o tysiącu sprawach.
And discuss.2PL about LOC thousand LOC things LOC
‘And we are talking about a thousand things.’

In the 1800m word corpus, there are 15 hits with INSTRUMENTAL on the quantified noun, 17 hits with DATIVE, and 139 hits with LOCATIVE (unverified hits).

Together, this suggests a grammaticalization cline, which 1000 is participating in:

(68) Noun > Deficient noun > Numeral

Conclusion: Numeral 1000 is optionally a lexical noun, which appears in the quantity domain of the quantified noun. Its noun-like properties derive from its noun-like functional structure.

5. Case study #2: English quantificational nouns

(69) A lot of books
(70) A ton of books
(71) A bunch of books

5.1 Properties of Q-nouns

5.1.1 Q-nouns are semi-lexical quantifiers

English numerals and quantifiers can combine directly with the noun they modify, though some require an indefinite article. They indicate quantity.

(72) Many/some/three books
(73) A few/a hundred/a dozen books

Q-noun pseudopartitives share this quantifying function, despite also appearing with of.

(74) A lot / lots / a ton / tons / a bunch of books

Together, this shows that quantifiers in English can have a variety of forms:

(75) Morphological expression of the quantifying function:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>lot / ton / bunch / number</td>
<td>of</td>
<td>books</td>
</tr>
<tr>
<td>a</td>
<td>few / hundred / dozen</td>
<td></td>
<td>books</td>
</tr>
<tr>
<td></td>
<td>many / some / three</td>
<td></td>
<td>books</td>
</tr>
</tbody>
</table>
Hypothesis: English Q-nouns are semi-lexical roots, surfacing in the quantity domain. Their roots contribute to the differences in their degree of ‘many-ness’ (in addition to a Q-feature which places them in the quantity domain of the quantified noun).

(76) \[ \text{ton} \succ \text{lot} \succ \text{bunch} \] \hspace{1cm} \text{Degree of many-ness}

5.1.2 Q-noun idiosyncrasies

The quantificational nouns (Q-nouns), \textit{lot}, \textit{ton}, and \textit{bunch}, cannot function as agreement targets, despite appearing to be singular nouns:

(77) A lot / lots of students weren’t studying.
(78) A ton / tons of students weren’t studying.
(79) A bunch of students weren’t studying.

They also show minor idiosyncratic differences; I report on plurality and definiteness.

Plurality: \textit{lot} and \textit{ton} can be morphologically plural, \textit{bunch} not.

\textit{Lot, ton}

(80) A lot of people
(81) Lots of people

(82) A ton of mistakes \hspace{1cm} \text{Compare:} A ton of bricks (=2000 lbs. of bricks)
(83) Tons of mistakes \hspace{1cm} \text{Compare:} Two tons of bricks (=4000 lbs. of bricks)

This plurality does not seem to indicate semantic plurality. Notice also that numerals and quantifiers cannot quantify these plural forms:

(84) *Two/many/a few lots of people
(85) *One lot of people \hspace{1cm} \text{\textit{/might be acceptable under a collection reading}}
(86) #Two / many / a few tons of mistakes
(87) # One ton of mistakes

\textit{Bunch}

(88) *Bunches of people/books/stuff

(89) Competing “literal” bunch interpretation:
a. She hung more crystals and \textit{bunches of herbs} throughout the room. (COCA: LiteraryRev 1990)
b. Angelo purchased two \textit{bunches of cut flowers}. (COCA: Bk:Blindsight 1992)

The lack of plurality seems to be subject to variation:

(90) I don’t have \textit{bunches of time} left. (COCA: CBS_Morning 1992)
(91) Well, I’ve had \textit{bunches of different kinds of memory loss}. (COCA: NPR 2013)
**Definiteness:** *Ton* can appear with DP material, *lot* and *bunch* cannot.

**Ton**

(92) Definite determiner (*ton/tons*):
   a. I quite frankly was very surprised that despite the *ton of hype* and all the excitement over Howard’s debut, the ratings weren’t much bigger than they were. (COCA: CNN_Showbiz 2012)
   b. Had I looked into a crystal ball and seen the *tons of extra stress* this would cause, well, frustration is getting the upper hand now. (COCA: AssocPress 2007)

(93) Demonstrative (*ton/tons*):
   a. I don’t know who helped this fellow take out this *ton of junk food*. (COCA: SatEvenPost 2008)
   b. the “hundreds of thousands of people” who wrote to CBS and all those *tons of nuts*… (COCA: WashPost 2007)

**Lot**

(94) *The lot of people
(95) *This/that lot of people

(96) *The lots of butter
(97) *These/those lots of butter

(98) Competing interpretation - collection *lot*:
   a. Old maids, the *lot of you* (COCA: Bk:IrishBlood 2015)
   b. I would prefer to damn this *lot of unsufferable bores* to the netherworld (COCA: Bk:MyLordVampire 2012)

**Bunch:**

(99) *The bunch of people
(100) *This/that bunch of people

(101) Competing interpretation - literal *bunch*:
   a. He plucks a dark-blue fruit from the *bunch of grapes* he is holding and puts it up to his lips. (COCA: Horticulture 1990)
   b. The *bunch of keys* that felt so heavy in my jacket pocket was totally unnecessary. (COCA: ChicagoRev 2002)

Literal *bunch* is syntactically singular, while Q-noun *bunch* is not. We can use the phrase *one-by-one*, which needs plurality, to exclude literal *bunch*:

(102) #He ate the bunch of grapes one-by-one.
(103) #He dropped the bunch of keys into the water one-by-one.
(104) #He ate that / one bunch of grapes one-by-one.
(105) #He dropped that / one bunch of keys in the water one-by-one.
These idiosyncrasies can be captured in the lexical specifications of the Q-noun (they do not seem to be properties of the pseudopartitive construction itself):

\[(106)\]  
Lot: \([\sqrt{lot}, \text{Q}, \text{INDEF}]\) //must occur in indefinite contexts
Ton: \([\sqrt{ton}, \text{Q}]\)
Bunch: \([\sqrt{bunch}, \text{Q}, \text{INDEF}, \neg\#_{\text{pl}}]\) //must occur in indefinite contexts; no plural

5.2 Analyzing Q-noun pseudopartitives

5.2.1 Structure of a Q-noun pseudopartitive

Pseudopartitives in many other languages look similar to the Q-noun pseudopartitives, differing only in the presence of a particle of:

\[(107)\] Een hoop mensen Dutch
‘A lot of people’

For such languages, the Q-noun has been analyzed as the head of a projection in the functional structure of the N2, e.g. QP (Löbel 1989), #P (Grestenberger 2015), ClP/MP (Stavrou 2003), or nP (Hankamer and Mikkelsen 2008).

\[(108)\] Juxtaposition construction (pseudopartitives)

\[
\begin{array}{c}
\text{DP} \\
\text{D} \\
\text{XP} (= \text{QP, #P, ClP, MP, nP}) \\
\text{Q-noun} \\
\text{N2}
\end{array}
\]

Analyses of languages which need a mediating particle like English, often try to place the of in the structure. Stickney (2004, 2009), for example, assumes a meaningless Functional Projection (FP) to host of.


\[
\begin{array}{c}
\text{DP} \\
\text{D} \\
a \\
\text{MP} \\
\text{M} \\
bunch \text{FP} \\
of \\
\text{NP}
\end{array}
\]
Ignoring of (for now), I assume a juxtaposition structure, labeling it QP. I take the Q-noun to be a root daughter of Q.

(110) Q-noun pseudopartitive structure

\[
\text{DP} \\
\text{D} \quad \text{QP} \\
\text{Q} \quad \sqrt{P} \\
\sqrt{\text{Q-noun}} \quad \text{N2}
\]

### 5.2.2 The indefinite article

Despite the indefinite article, the construction is not singular.

(111) Pseudopartitive Q-nouns:
   a. A lot of **people were/’was** invited to the party.
   b. A ton of **people were/’was** invited to the party.
   c. A bunch of **people were/’was** invited to the party.

This is also true for other indefinites in quantifying expressions:

*Quantifiers which need an indefinite article*

(112) A hundred / a couple people **were** stalking the dragon.
(113) A few people **were** fighting in the ballroom.
(114) A dozen eggs **were** smashed on Paul’s head.

*Modified many (obligatory indefinite)*

(115) A great/ good many donkeys **have** fallen to the butcher’s knife.

*Modified lower numerals (obligatory indefinite)*

(116) **A mere eight companies own** nearly all of mainstream media journalism. (COCA: USAToday 2007)
(117) **A reported 4,000 delegates were** among the 10,000 conventioneers gathered in Charlotte, NC, for the six-day confab. (COCA: Jet 1996)
(118) **A suffocating 92,000 fans were** in the seats for each game at the L.A. Memorial Coliseum. (COCA: Chicago 2005)

*Modified plural higher numerals (obligatory indefinite)*

(119) Bypassing or tampering with power supplier meters is a growing problem, costing **an estimated hundreds of billions of dollars** worldwide and accounting for 10% to 40% of all energy use in various countries. (COCA: Futurist 2013)
Proposal: The indefinite is a “dummy” article which is used as a last resort strategy to fill an empty Q-head.

- Lyons (1999): *A(n) does not mark indefiniteness, but cardinality. It is a (singular) default cardinality marker when no numeral, quantifier, or plurality occurs.
- Based on the examples above, I take *a(n) to be unspecified for number.
- Thus, the use of a default cardinality marker with a Q-noun indicates the presence of an empty QP.
- The Q-noun is a root which cannot fill Q- the default cardinality marker is invoked.

(120) Q-noun pseudopartitive structure

What about apparent plural lots and tons?

Proposal: By the analysis above, -*s must sit in Q. As a bound morpheme, it attracts the root. (What does -*s add to the interpretation? Intensification of sorts?)

(121) Plurality in the pseudopartitive (lots, tons)

Implication: The presence or absence of *a(n) may indicate the level of grammaticalization the quantifier has undergone. More grammaticalized quantifiers (numerals, many) can lexicalize Q, while less grammaticalized quantifiers (Q-nouns) cannot.
5.2.3 The particle of

Proposal: *Of* is a marker of nominality, when multiple nominals are present in a single domain (cf. dependent case, Marantz 1991, Baker 2015, Levin 2015, a.o.). The presence of two roots seems to be enough (but perhaps there is more to the structure) to trigger *of*.

(122) 
```
  DP
   D QP
      Q √P
        a √Q-noun #P
          of # √N2
```

If *of* is a morphological marker, it explains why Dutch might lack such a particle – it lacks this particular mechanism of marking nominals in a single domain.

(123) Een hoop (*van) mensen 
A lot (*of) people
‘A lot of people’

Implication: The presence or absence of *of* may also indicate the level of grammaticalization, but in this case, whether the quantifier is noun-enough to trigger *of*.

This helps us understand the table below, and why we don’t find quantifiers with *of* but without a default cardinality marker:

(124) Morphological expression of the quantifying function:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>lot / ton / bunch / number</td>
<td>of</td>
<td>books</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>few / hundred / dozen</td>
<td></td>
<td>books</td>
<td></td>
</tr>
<tr>
<td></td>
<td>many / some / three</td>
<td></td>
<td>books</td>
<td></td>
</tr>
</tbody>
</table>

- *A(n):* Can the quantifier lexicalize Q?
- *Of:* Is the quantifier still “noun-enough” to need differentiation from the noun?
5.2.4 Agreement

Q-nouns surface as a bare root under a QP, and carry no number of their own:

\[
\begin{array}{c}
\text{QP} \\
\text{Q} \quad \sqrt{\text{P}} \\
\sqrt{\text{Q-noun}} \\
\text{a/(-s)} \\
\ldots
\end{array}
\]

Lacking number, they cannot control agreement or other processes sensitive to number.

**Conclusion:** Q-nouns are quantifiers (with roots), which resemble lexical nouns. This resemblance arises because they are not grammaticalized enough as quantifiers to lexicalize the Q-head or alleviate the need for *of*, unlike numerals or other quantifiers in English.

6. Conclusions

- We seem to find a similarity between lexical categories and quantificational expressions in language.
- Two case studies were considered: Polish numerals (numeral 1000), and English Q-nouns *lot*, *ton*, and *bunch*
- Numeral 1000 was analyzed as having much of the same functional structure as a lexical noun (presumably due to historical reasons), despite appearing in the quantity domain of the noun. This meant that it had a number of noun-like properties, despite functioning as a quantifier.
- Q-nouns *bunch*, *lot*, and *ton* were analyzed as roots in the quantity domain. Their similarity to nouns arose from (a) properties of Q (must be filled), (b) their inability to lexicalize Q, (c) properties of *of*, and (d) their more lexical nature than other quantifiers.

7. References


H. van Riemsdijk (Eds.), *Sounds of Silence: Empty Elements in Syntax and Phonology* (pp. 105-144). Amsterdam: Elsevier.


