Universal semantic features and the typology of numerals

Introduction. The paper proposes a unified morpho-semantic account for the typological variation in form and meaning of cardinal numerals. In particular, we investigate the morphological marking of different types of cardinals and argue that despite an apparent morphological chaos, it is possible to identify cross-linguistically stable semantic ingredients, which compositionally provide the attested types of numerals. We adopt the framework of Nanosyntax (Starke 2009 et seq.) as a model of morphology which, when applied to the semantic structures we propose, delivers the relevant marking patterns. The model we develop is broadly based on the idea that the meaning components are uniformly structured across languages, and they must all be pronounced, though languages differ in how they pronounce them.

The asymmetry. Cardinals can have different functions including what we will refer to as ABSTRACT COUNTING, i.e., reference to a number concept, and OBJECT COUNTING, i.e., quantification over individuals (Bultinck 2005, Rothstein 2017). Interestingly, languages often distinguish formally between the two flavors (Hurford 1998). For instance, in Japanese a form used to refer to mathematical entities, see (1-a), differs from the one conveying the cardinality of a particular set of objects in (1-b) (Sudo 2016). Though both expressions contain a common core, e.g., *yon*, the object-counting function requires an additional morpheme, e.g., *rin*, usually referred to as a classifier. Such an asymmetry is a cross-linguistically relatively frequent pattern which suggests that the abstract-counting function is basic whereas the object-counting function is derived from it both morphologically and semantically.

(1) a. ni tasu ni-wa **yon-(*ko)**-da
two plus two-TOP four-CL-COP
'Two plus two is four.'

b. **yon-*(rin)**-no hana
four-CL-GEN flower
'four flowers'

Symmetric languages. In a number of languages, however, we observe no such asymmetry as in (1). For instance, in English both functions are expressed by the same formal exponent, see (2), suggesting that the bare numeral itself incorporates a classifier semantics (Krifka 1995). In other words, the form of *four* is ambiguous. In one use, it is semantically equivalent to *yon-rin*, in another use to *yon*.

(2) a. Two plus two is **four**.

b. **four** roses

Inverse languages. The most intriguing morphological facts come from Arabic. In this language, abstract counting is expressed by a morphologically more complex form than object counting, see (3) (Fassi Fehri 2018). This pattern seemingly implicates a reverse asymmetry, i.e., that compared to the object-counting function the abstract-counting function has some extra meaning which needs to be introduced by an additional morpheme, e.g., a gender marker. However, admitting this would jeopardize a morpho-semantic explanation of the widespread asymmetry illustrated in (1) as well as any unified typology of numerals. What we need to capture is the fact that Arabic exists, but that it is rare.

(3) a. **ṯaalaṯ-*(at)**-un tusawii ʔiṯnayni zaʔid waaḥid
three-FEM-NOM equals two plus one
'Three equals two plus one.'

b. **ṯaalaṯ-(*at)**-u banaatin
three-FEM-NOM girls
'three girls'

Universal semantic features. In order to account for the data, we propose the ingredients in (4) to be part of the universal underlying structure of numerals. We assume three syntactic heads and the standard type-shifting operation iota which can lower the type of a predicate (Partee 1987). NUM is a lower bounded set, e.g., $\{4, 5, 6, ...\}$. CARD takes a set of numbers as its input and returns a set of entities

(Landman 2003). Its goal is to form an expression that can be used for counting actual objects. Finally, EXACT is interpreted as a predicate modifier which yields a subset consisting of the smallest elements of a set of numbers or entities. As such, it introduces the upper bound, and thus contributes exactness to the meaning of a numeral. Due to its typal flexibility it can attach either to [NUM] or to [CARD NUM].

(4) a. $[\![\text{NUM}]\!]_{\langle n,t\rangle} = \lambda m_n.m \geq n$ c. $[\![\text{CARD}]\!]_{\langle\langle n,t\rangle,\langle e,t\rangle\rangle} = \lambda P_{\langle n,t\rangle}\lambda x_e.\#(x) \in P$
b. iota = $\lambda P_{\langle \alpha,t\rangle}.\iota x_\alpha[P(x)]$ d. $[\![\text{EXACT}]\!]_{\langle\langle \alpha,t\rangle,\langle \alpha,t\rangle\rangle} = \lambda P_{\langle \alpha,t\rangle}.\text{MIN}(P)$

Composition. Combining these ingredients in a compositional fashion leads to the semantic structures in (5). Application of iota to (5-a) returns the unique integer equal to the relevant number, e.g., 4 (Kennedy 2015). A resulting expression is of type n and can be used as a name of a number concept. On the other hand, (5-b)–(5-d) are object-counting expressions. Note that both (5-c) and (5-d) correctly predict obligatory exact interpretations of numerals in their predicative uses (Partee 1987, Geurts 2006). Moreover, they do not exclude unilaterally bounded interpretations of numerals in attributive position due to existential closure over the individual variable. Hence, for the bilaterally bounded interpretation we assume an exhaustivity operator which composes with the meaning of the whole sentence (Chierchia 2004).

(5) a. [EXACT NUM] ABSTR.COUNT c. [EXACT [CARD NUM]] OBJ.COUNT
b. [CARD NUM] OBJ.COUNT d. [CARD [EXACT NUM]] OBJ.COUNT

The non-terminal lexicalization model. To account for the morphological patterns, we adopt the view that lexical entries link morphemes to potentially complex syntactic/semantic structures. In addition, following Starke (2009), we assume that the Superset Principle allows a given morpheme to pronounce *any sub-constituent* contained in its lexical entry. For instance, a lexical entry such as (6-a) can also as pronounce the structure (6-b) or (6-c) since both of these structures are its sub-constituents. However, it cannot pronounce a structure like (6-d), since (6-d) is not a sub-constituent of (6-a). In addition, we assume that there are no cardinals pronouncing only [NUM], but we will independently justify its relevance based on the semantics and suppletive morphology of ordinals and multipliers.

(6) a. [EXACT [CARD NUM]] b. [CARD NUM] c. [NUM] d. [EXACT NUM]

Typology. This system is able to derive all the attested variation by treating different types of numerals as lexicalizations of different structures derived from the universal semantic components, see Table below. In symmetric languages, numerals are stored as complete structures pronouncing all the three heads, which allows them to cover both the abstract-counting and the object-counting function, including unilaterally and bilaterally bounded readings. On the other hand, in asymmetric languages numerals lexicalize only the abstract-counting meaning, and thus require additional morphology in order to be able to modify a noun or be used as a predicate. Finally, in inverse languages, object counting numerals are specified with the reversed hierarchy of EXACT and CARD, i.e., as [EXACT [CARD NUM]. With such an entry, the numeral can function as an object-counting expression. However, it cannot function as an abstract-counting numeral. This is because the structure [EXACT NUM] is not a sub-constituent of its lexical entry. Therefore, the element EXACT has to be expressed separately. In addition, the system also allows to account for split languages such as Chol which exhibits simultaneously both the asymmetric and symmetric system though in different numerals (Bale & Coon 2014) and suppletive forms of abstract- and object-counting numerals within a single language, e.g., Maltese *tnejn* ~ *żewg* for 2 (Hurford 1998).

TYPE	LANGUAGE	NUMERAL	ADDITIONAL MORPHEME
SYMMETRIC	English	*four* ⇔ [CARD [EXACT NUM]]	Ø
ASYMMETRIC	Japanese	*yon* ⇔ [EXACT NUM]	*rin* ⇔ [CARD]
INVERSE	Arabic	*ṯaalaṯ* ⇔ [EXACT [CARD NUM]]	*at* ⇔ [EXACT]

References. Bale & Coon (2014) *Classifiers are for numerals, not for nouns* • Bultinck (2005) *Numerous meanings* • Chierchia (2004) *Scalar implicatures, polarity phenomena and the syntax/ pragmatics interface* • Fassi Fehri (2018) *Constructing feminine to mean* • Hurford (1998) *The interaction between numerals and nouns* • Kennedy (2015) *A 'de-Fregean' semantics for modified and unmodified numerals* • Krifka (1995) *Common nouns: A contrastive analysis of Chinese and English* • Landman (2003) *Predicate argument mismatches and the adjectival theory of indefinites* • Partee (1987) *Noun phrase interpretation and type-shifting principles* • Rothstein (2017) *Semantics for counting and measuring* • Starke (2009) *Nanosyntax: A short primer to a new approach to language* • Sudo (2016) *The Semantic role of classifiers in Japanese*