Singulatives in Hasawi Arabic and Welsh

Luca Molinari, Ruoying Zhao, Marcin Wagiel, Abdullah Almuhaysh, Victoria Noble

Intro & aim: "Singulatives" (SGTV) are morphologically marked singulars derived from an unmarked base generally referred to as "collective" (COLL), typically denoting fruit/vegetables, small animals etc. that usually appear in groups. However, the properties of COLL and SGTV vary across languages and thus call for different analyses. The present contribution aims to provide an account for these different paradigms, shedding light on different morpho-syntactic ways of encoding COLL/SGTV cross-linguistically.

Morphological paradigms: COLL behave like regular plurals (PL) in some languages: in **Welsh** (**W**; Nurmio 2017), for instance, COLL trigger PL agreement on verbs and require PL anaphoric pronouns (1). On the other hand, **Hasawi Arabic** (**HA**, eastern Saudi Arabia - fieldwork data) COLL behave like singular (SG) terms (2). COLL in HA are always MASC, while the derived SGTV is FEM and can be further pluralized (3).[1] In W, there are two SGTV suffixes which match the Gender of the COLL (4-a)-(4-b), but neither can be pluralized.

(1) yn rhwydo *pysgod* ddydd ar ôl dydd, yn **eu** glanhau a **'u** gwerthu
PROG catch.VERB-NOUN *fish*.**COLL.M** day after day PROG **their** clean.V-N and **their** sell.V-N
'catching *fish* day after day, cleaning them and selling them' (W; Nurmio 2017: p. 2)

(2) ʔəs-*səmak* **ji**-ʕiiʃ fi-l-maaj. (HA)
DEF-*fish*.**COLL.M** **3SG**-live in-DEF-water
'*Fish* live in water.'

(3) səmak ~ smək-**ah** ~ smək-**aat**
fish.COLL.M fish-**SGTV.F** fish-**SGTV.F.PL**
'fish (mass) ~ a fish ~ (multiple) fish' (HA)

(4) a. pysgod ~ pysgod-**yn**
fish.COLL.M fish-**SGTV.M**
'fish(es) ~ a fish'

b. sêr/sŷr ~ ser-**en**
star.COLL.F star-**SGTV.F**
'stars ~ a star' (W)

HA pattern: For countable MASC animate nouns, the SGTV suffix *-at* yields the FEM counterpart, like in other Arabic dialects (e.g., *malək* 'king' > *malək-ah* 'queen'; see Fassi Fehri 2018). This suffix also attaches to COLL, which have kind reference and no access to natural units in their denotation. We distinguish **two patterns: ① objects** and **② substances**. For ①, *-at* returns a set of **natural units** of that to enable counting (θ*alaa*θ *smək-aat* 'three fish'). For ②, *-at* enables counting by sorting the substance into **integrated portions** (pieces, etc.) or, in certain cases, **(sub)types**.

HA Number paradigm gaps: HA marks the DUAL (***-een***). PL marking for the vast majority of words takes the form of **stem-internal vowel lengthening** ('**broken plural**', BRKN PL). For patterns ① and ②, the BRKN PL and the **direct DUAL (i.e. no intervening SGTV)** are used for counting types ((10)c-d). Some HA substance nouns lack either the BRKN PL or the direct DUAL in the paradigm (5), in which case the SGTV DUAL (***at-een***) and **plural of the singulative** (PL OF SGTV) (marked by ***-aat***, i.e., ***-at* + vowel lengthening**) allow counting of types in addition to natural units/portions.

(5) *keek* ~ *keek-ah* ~ *keek-at-een* ~ *keek-aat*
cake.M.SG cake-SGTV.F cake-SGTV.F-DUAL cake-SGTV.F.PL
'cake (mass)' ~'one cake/**a type** of cake' ~'two cakes/**two types** of cake' ~'cakes/**types** of cake'

Analysis of HA: The account of SGTV in Kipsigis proposed by Blix (2022) is not appropriate for HA, given that both the mass-like properties of COLL and the fact that the SGTV itself can be pluralized (but see below for W). We observe that: (i) counting by natural units requires explicit SGTV marking, and (ii) COLL can be directly counted by subtype, unless the morphology is missing in the paradigm ((5)-(10)g-i), in which case SGTV is required. We thus take **countability to have a structural correlate, i.e., an fseq above NP**. The NP enters the derivation as a mass-like noun (as in Borer 2005) which gets partitioned either by type via a TYPEP projection or by natural units/portions of matter via UNITP (stacked on top of TYPEP). Building on this assumption, we model different types of nouns according to their different structural sizes, as in (6) (subsuming referential projections under NP and leaving out here Gender and Number).

(6) a. count nouns: [UNITP [TYPEP [NP]]] *malək* 'king'
b. COLL/substance-mass nouns allowing direct dual/BRKN PL: [TYPEP [NP]] *səmak* 'fish'
c. substance-mass nouns lacking direct dual/BRKN PL: [NP] *keek* 'cake'

The lexical entry for the SGTV suffix *-at* is shown in (7-c). It contains a pointer to a *caboodle* item (a mechanism introduced by Caha 2022) given in (7-b) which in turn contains a pointer to a further *caboodle* item (7-a). We also specify a lexical entry for the DUAL (8) and for PL (9).

[1] In HA, there is a **morphophonological rule** which **turns the *-t* into an *-h*** if *-at* is the final morpheme of a word.

(7) a. /*caboodle*$_2$/ ⇔ [UnitP [TypeP]]
b. /*caboodle*$_1$/ ⇔ [FemP [MascP [GendP [→ *caboodle*$_2$]]]]
c. /*-at*/ ⇔ [SgP [→ *caboodle*$_1$]]

(8) /*-een*/ ⇔ [DualP]

(9) /LENGTHENING/ ⇔ [PlP [DualP]]

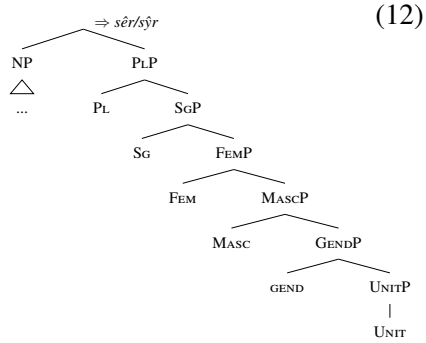
Some examples of lexicalizations are given in (10) for different noun types: **count (a-b)**, **object coll (c-f)**, **substance-mass (g-i)**. UnitP is spelled out by default together with the NP in countable masc nouns ((10)a), but is introduced by *-at* in countable fem ones, forcing backtracking down to NP ((10)b) (indeed, there is no fem coll base). Object coll ((10)c) spell out by default TypeP with their NP (so direct dual/brkn pl, if available, counts types). They can be singulativized by projecting UnitP, lexicalized by *-at* which again forces backtracking to NP. Words like 'cake' ((10)g-i) only spell out NP: this explains the unavailability of direct dual and brkn pl, both of which need TypeP. For this reason, *-at* fills in and enable counting by either types (realizing TypeP only, which still lexicalizes *caboodle2*) or integrated portions (realizing [UnitP [TypeP]]). The complexity of a lexical entry with two pointers explains why **sgtv counting types are very rare across languages** (and not acknowledged, see Dali & Mathieu 2021: p. 276). The pl's lexical entry overwrites dual marking (9) and is spelled out by a non-segmental morpheme causing **lengthening of the closest constituent**. Therefore, for brkn pl (counting types), lengthening applies to the nominal root (yielding *əsmaak* in (10)d),[2] and for pl of sgtv (counting units) it applies to *-at* (yielding *smək-aat* in (10)f).

(10)

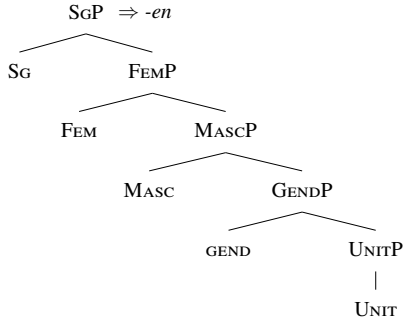
Hasawi Arabic	NP	TypeP	UnitP	GendP	MascP	FemP	SgP	DualP	PlP
a. '(one/two) king(s)' (sg/dual)				malək				een	
b. '(one/two) queen(s)' (sg/dual)	malək			at				een	
c. '(two types of) fish' (object coll/dual)		səmak						een	
d. '(types of) fish' (brkn pl)		səmak						LENGTHENING	
e. '(one/two) fish' (sgtv/dual)	səmak			at				een	
f. '(one/many) fish' (pl of sgtv)	səmak			at				LENGTHENING	
g. 'cake' (coll/subst.mass)	keek								
h. 'one/two type(s)/piece(s) of cake' (sgtv/dual)	keek			at				een	
i. 'one/many type(s)/piece(s) of cake' (pl of sgtv)	keek			at				LENGTHENING	

Analysis of W: Since coll in W are shown to pattern with regular pl (Nurmio 2017) wrt morphosyntactic properties, and given the impossibility of pluralizing the obtained sgtv, we extend the analysis by Blix (2022) to W as well. We contend that W coll are stored as blixemic entries featuring a pl feature (11) (thus deriving the impossibility of pluralizing the derived sgtv). The sg form *-en*, not being a sub-tree of coll, forces the insertion of the sgtv suffix (12) (the masc counterparts *pysgod* and *-yn* (4-a) lack FemP).

(11)



(12)



Conclusion: This paper derives the differences in the morphosyntactic properties of the coll/sgtv paradigm in W and HA from the stored shape of the lexical entry of coll. It extends Blix's analysis to W and provides a **novel nanosyntactic analysis** of the sgtv suffix *-at* in HA, accounting for the complex marking pattern found in this language and the unit vs. (sub)type readings of *-at*. The analysis of HA further allows an interesting cross-linguistic parallel between *-at* and the singulative suffix *-in-* in Slavic which precedes the fem.sg marker and can only partition by natural units (Wągiel & Shlikhutka 2023): Ukrainian *popil* (ash.m.sg) 'ash' ~ *popel-**yn**-a* (ash-**sgtv**-f.sg) 'a speck of ash' ~ *popel-**yn**-y* (ash-**sgtv**-pl) 'specks of ash'. These facts can be captured if *-in-* lexicalizes the entry in (7-b) but without the pointer (i.e., *-in-* ⇔ [FemP [MascP [GendP [UnitP [TypeP]]]]]).

References: Blix, H. (2022). Interface legibility and nominal classification: a nanosyntactic account of Kipsigis singulatives. *Glossa*. • Borer, H. (2005). *In Name only: Structuring sense*. • Caha, P. (2022). The marking of mass, count and plural denotations in multidimensional paradigms. *Studia Linguistica*. • Dali, M. & Mathieu, E. (2021). Singulative systems. In Cabredo Hofherr, P. & J. Doetjes (eds.), *The Oxford Handbook of Grammatical Number*. • Fassi Fehri, A. (2018). *Constructing Feminine to Mean: Gender, Number, Numeral, and Quantifier Extensions in Arabic* • Nurmio, S. (2017). Collective Nouns in Welsh: A Noun Category

[2]For certain borrowed words with non-native roots, the phonological rule for broken plural is not applicable, so HA employs the suffix *-aat*, e.g., *kombjuutaɪ* 'computer' (count) ~ *kombjuutɪ-**aat*** 'computers'.

or a Plural Allomorph? *Transactions of the Philological Society*. • Wągiel, M. & Shlikhutka, N. (2023). Clustering and declustering things: The meaning of collective and singulative morphology in Ukrainian. *Proceedings of SALT33*.