Category gradience in a feature-based generative approach: English pseudopartitives

Introduction: There is a tension between formal generative approaches which assume a discrete feature-based approach to categories (e.g. Zeijlstra 2023) and cognitive-functional approaches which view categories in terms of gradience or prototypes (e.g. Keizer 2023). Within the latter approach, the notion of gradience has been linked to patterns of grammaticalization (Traugott & Trousdale 2010), where gradience involves the synchronic layering of ongoing processes of change (e.g. Brems 2011). In the generative tradition, similar claims have been made for semi-lexicality (Corver & van Riemsdijk 2001), e.g. Cavirani-Pots (2020), who redefines verbal semi-lexicality in terms of a grammaticalization cline. In this talk, I explore how the gradience posited for pseudopartitives in cognitive-functional approaches (Keizer 2007, Brems 2011, Ten Wolde 2023) is accounted for in a feature-based generative approach, building on the idea that gradience is linked to grammaticalization.

Pseudopartitives: In English, the pseudopartitive is a binominal construction of the form "N1 of N2", in which the N2 is a substance noun and the N1 is a semi-lexical or functional noun which performs some function on the N2, such as quantification (*a lot of books*) or containment (*a bottle of wine*). Depending on its use, the N1 in the binominal may be more or less prototypical for a noun, suggesting gradiency within a single lexical item. Brems (2011), for example, investigates size nouns and identifies a range of uses with corresponding morphosyntactic properties. The size noun *heap*, for instance, has a literal use (*heap* as a haphazard pile of objects) (1), a quantifier use (2), and a valuing quantifier use, which implies negative semantic prosody (3).

(1) Each home […] was reduced to a heap of rubble (Brems 2011: 135, ex. 4.19)
(2) They went through my bags, searched me and asked a heap of questions (ibid: 140, ex. 4.27)
(3) […] he's been talking a heap of shit to me […] (ibid: 146, ex.4.57)

In Brems' account, there is delexicalization of the N1, allowing greater collocational flexibility, which eventually cumulates in a reanalysis of the head-modifier construction (e.g. (1)) to a modifier-head construction (e.g. (2) and (3)). Each construction is simultaneously available to the speaker, representing a synchronic layering of the different grammaticalization steps from noun to quantifier.

The present study: I report on the results of a corpus study of metaphorical N1s in English. Metaphorical N1s were chosen to give a better understanding of the early stages of grammaticalization in combination with our current understanding of established N1s as shown in e.g. Keizer (2007) and Brems (2011). Exhaustive searches for the lexemes *pond*, *geyser*, *ocean* and *waterfall* followed by an *of* were conducted in the Corpus of Contemporary American English (Davies 2008-), hereafter "COCA". Examples where the N2 contained high functional material (determiners, quantifiers) or constituted a pronoun or clause were excluded, in line with the characterization of pseudopartitives in Alexiadou et al (2007). The remaining examples (N=1304) were coded for determiner type on N1, pre-modifiers on N1 and N2, subject-verb agreement, etc. *Ocean* was the most frequent type (N=951), followed by *geyser* (N=136), *waterfall* (N=123), and *pond* (N=94).

Head-modifier uses: All lexemes allow a construction in which the N2 identifies the contents of the N1. This takes the form of a lexical use where the N2 denotes water (4), but also a delexicalized use, where the N2 denotes a non-canonical substance (5); these are literal uses of the N1. Metaphorical uses also occur, where the N2 is an imagined substance of a metaphorical N1 (6). Metaphorical uses may facilitate an eventual use as a quantifier, with speakers reanalyzing the N1 as quantificational, particularly if the N1 has a lexical size feature to foreground (e.g. *ocean*). This may be the case with examples like (7), which seems compatible with both a metaphorical and quantificational reading.

(4) a pond of mucky water; geysers of boiling water; a vast ocean of liquid water; the waterfall of cool and cleansing water (COCA)
(5) ponds of liquid tar; geysers of natural gas; an ocean of magma; waterfalls of slime (COCA)
(6) dipping my toes into the pond of blogging; Raymie is a geyser of gossip and hard news (COCA)
(7) An ocean of asphalt greets visitors. (COCA)

Such uses instantiate a head-classifier binominal, in the terminology of Ten Wolde (2023), which is a head-modifier use. A single syntactic structure can be assumed, where the N1 denotes a lexical noun which takes an *of*-PP as its complement. Whether it is interpreted literally or figuratively is determined pragmatically, i.e. as a function of the N1, N2, and context. This analysis predicts that

verbal agreement targets N1, that pre-modifiers target the full binominal, and that N1 is free to combine with any type of determiner. These predictions appear to be true in the dataset.

Modifier-head uses: Brems (2011) identifies quantifying and valuing quantifying modifier-head uses, which is expanded by Ten Wolde (2023) to include evaluative binominals (e.g. *the ocean of a bed*, COCA); among these types, only quantifying uses were frequent enough for analysis, and largely only for *ocean* (8). For this reason, the remaining analysis focuses on the quantifying use of *ocean*.

(8) a. Belgium brews an ocean of beer – 350 kinds (COCA)
b. As doctors deal with an ocean of paperwork […], patients suffer. (COCA)

Verbal agreement showed agreement with both N1 and N2 (9). This contrasts with established quantifier N1s like *lot*, which require N2 agreement (Keizer 2007). All types of determiners were found (10), suggesting that there are no determiner restrictions; this is again in contrast to *lot*, which Keizer (2007) and Brems (2011) report to prefer indefinite contexts. Of the pre-modifiers found, they targeted the lexical semantics of the N1 (*deep, endless, vast*) (11a), the quantifier meaning (*whole*) (11b) or the binominal (11c). This is another difference from the previous literature, where Brems (2011) found that quantifier uses combined with quantifier-reinforcing pre-modifiers.

(9) a. An ocean of men were swarming toward them. (COCA)
b. An ocean of questions surfaces in the wake of his essay. (COCA)

(10) a. Oceans of free time; the oceans of money pouring in (COCA)
b. an ocean of radio waves; the ocean of Bosnian atrocities (COCA)

(11) a. a vast ocean of grammatical phenomena b. a whole ocean of pretty Australian puppies
c. this hidden ocean of funds (COCA)

Analysis: I propose micro-steps in the grammaticalization from noun to quantifier, generalizing and expanding on the grammaticalization cline of Cavirani-Pots (2020, 2022). At Stage 0, the N1 is fully lexical. This means it contains a root embedded under a categorizer *n* which occurs in a canonical nominal environment; the head-modifier use represents this stage. In Stages I and II, the N1 is semi-lexical; in Cavirani-Pots' system, this means a root is present, but it occurs in the functional domain of the N2. In Stage I semi-lexical, it first combines with a categorizer, whereas in Stage II semi-lexical it combines directly with a functional head. In Stage III, the N1 is functional (e.g. *many*) and does not contain a root; no *of* appears either, suggesting its presence correlates with the roothood of N1. I propose that verbal agreement indicates whether a lexical item is in Stage I or Stage II of semi-lexicality. In Stage I, the root is dominated by both *n* and Num, and therefore acts as an intervener for agreement with N2. In Stage II, the root combines directly with Q, not carrying number of its own. It no longer acts as an intervener, and agreement targets N2. The N1 *lot* has reached Stage II, hence why it never determines agreement; the N1 *ocean*, however, can realize both Stage I and Stage II, hence the varying agreement pattern. Finally, I propose that there may be a further step on the grammaticalization cline, Stage IIb, in which a definite DP is banned. This is the case for *lot* and fits with the argumentation of Alexiadou et al 2007 that true quantification occurs in indefinite contexts.

Conclusions: Gradience has taken a backseat in generative theorizing. Yet, gradience is important for understanding the structure of a category and how an element might shift into, out of, and within that category. In the present approach, changes in the position and dominating structure of the N1 lead to gradience, with the implication being that (sub-)categories are not continuous, but discrete.

References: Alexiadou, Haegeman, & Stavrou (2007). *Noun Phrase in the Generative perspective*. De Gruyter. ♦ Brems (2011). *Layering of size and type noun constructions in English*. De Gruyter. ♦ Cavirani-Pots (2020). *Roots in progress: Semi-lexicality in the Dutch and Afrikaans verbal domain*. LOT. ♦ Cavirani-Pots (2022). Roots in progress: A case study on Dutch semi-lexical verbs. Presentation at Penn State University. ♦ Corver & Van Riemsdijk (2001). *Semi-lexical categories: The function of content words and the content of function words*. De Gruyter. ♦ Davies (2008-). *The Corpus of Contemporary American English: 450 million words, 1990-present.* Available online at http://corpus.byu.edu/coca/. ♦ Keizer (2007). *The English Noun Phrase: The nature of linguistic categorization*. Cambridge UP. ♦ Keizer (2023). Word classes and gradience. In van Lier (Ed.) *The Oxford handbook of word classes.* Oxford UP. ♦ Ten Wolde (2023). *The English binominal noun phrase: A cognitive-functional approach*. Cambridge UP. ♦ Traugott & Trousdale (2010). *Gradience, gradualness, and grammaticalization*. John Benjamins. ♦ Zeilstra (2023). Word classes in Minimalist Syntax. In van Lier (Ed.), *The Oxford handbook of word classes*. Oxford UP.