Francesca Masini, Muriel Norde & Kristel Van Goethem

Approximation in morphology:
A state of the art

Abstract: This paper offers a state of the art of approximation within the larger domain of evaluative morphology. It provides an overview of the formal means employed by the morphology of different languages to express approximative meanings, as well as a survey of the specific approximative values that can be conveyed. We further discuss the (input and output) lexical categories involved in the expression of approximation crosslinguistically. On an intralinguistic level, we address the issue of competition between different strategies.

Keywords: evaluative morphology, approximation, vagueness, categorization, competition, cross-linguistic variation.

1. What is approximative morphology?*

1.1 Evaluative morphology and approximation

Evaluative morphology is by now a well-established domain of investigation (cf. e.g., Bauer 1997; Grandi & Körtvélyessy 2015; Körtvélyessy 2015). However, the semantic functions performed by evaluative morphemes have not been investigated in equal measure. Whereas functions such as diminution, augmentation and intensification have been studied quite extensively (cf., among many others, Dressler & Merlini Barbaresi 1994; Jurafsky 1996; Grandi 2002; Schneider 2003; Bakema & Geeraerts 2004; Prieto 2005; Körtvélyessy & Štekauer (eds.) 2011; Efthymiou 2015; Efthymiou, Fragaki & Markos 2015; Napoli 2017; Hendrikx 2019), others, like ‘approximation’, have received much less attention (see Amiot

* We are deeply grateful to the participants of the ApproxiMo – Approximation in morphology ‘discontinuous’ workshop, which was held online between December 2021 and May 2022. The workshop exceeded our expectations: it was a very successful event with rich interactions and stimulating discussions, which greatly contributed to the preparation of this introduction.
 APPROXIMATION IN MORPHOLOGY: A STATE OF THE ART

& Stosic 2022 for a similar argument). Consequently, we still know very little about how approximation works within morphology.

1.2 Previous research

Only very recently has approximative morphology started to attract the attention of scholars. The most studied approximative morpheme is definitely English -ish, which has undergone remarkable semantic and morphological developments in the course of time. In present-day English, approximative -ish expresses vagueness and speaker attitude, and it has been expanding its host classes from adjectives (e.g., warmish) to a variety of other categories (e.g., 9-ish, okay-ish) (e.g., Oltra-Massuet 2017; Kempf & Eitelmann 2018; Eitelmann, Haugland & Haumann 2020; Eitelmann & Haumann 2023). It has also been extensively discussed in the context of degrammaticalization because of its recent autonomous or “debonded” uses (Kuzmack 2007; Norde 2009), as in (1).

(1)    Ali: So, you feeling any better yet? - Me: Eh, ish ish [urbandictionary.com]

Other recent studies focused on the emergent approximative use of Italian simil- (e.g., simil-marsupio ‘sort of marsupium/pouch’, freddo simil siberiano ‘Siberian-like cold’) (Masini & Micheli 2020), and on the productivity, semantic profiles and categorical flexibility of a series of Dutch morphemes with ‘fake’ semantics (e.g., kunstgras ‘artificial grass’, namaak-wasabi ‘fake wasabi’, neppe cupcake ‘fake cupcake’) (Van Goethem & Norde 2020). A large-scale cross-linguistic study is Van Goethem, Norde & Masini (2021), which is a corpus analysis of pseudo- in eight European languages (Danish, Dutch, English, German, French, Italian, Spanish and Swedish), with the aim to compare their morphological and distributional properties (e.g., Italian pseudotifosi ‘pseudo-supporters’; Spanish pseudoartistas ‘pseudo-artists’; German pseudoreligiös ‘pseudo-religious’; Dutch pseudo-wetenschappen ‘pseudo-sciences’).

The fact that research into approximation is less established within morphological studies is also reflected by the abundance of terms that are used to refer to it, and by the lack of consensus on their use. For example, Grandi & Körtvélyessy (2015: 9, 11) use the triad ‘approximation/reduction/attenuation’ to refer to this functional domain. Other relevant terms are ‘deintensification’ (Körtvélyessy 2015: 63), ‘non-prototypicality’ (Cúneo 2015: 630), or ‘non-authenticity’, ‘fakeness’ and ‘imitation’ (Masini & Micheli 2020). Some of
these terms are typically associated with specific domains. For instance, ‘attenuation’ typically refers to ‘reduced degree of a quality’ in relation to adjectives (cf. e.g., Bauer 2002), like in French blanchâtre ‘whitish’. Whereas ‘attenuation’ is rather widespread, terms like ‘reduction’, ‘non-prototypicality’ or ‘deintensification’ appear to be less common.

For this Special Issue, we decided to use approximation as a cover term for the (complex) functional domain associated with all these values: among the many available terms, it seems to be the one broadest in scope and, at the same time, one of the least associated with a specific domain/phenomenon. Contrary to Rainer (2015: 1346), who includes ‘approximation/attenuation’ in the wider category of ‘intensification’ (which “comprises not only a high degree but all degrees of intensity”), we keep approximation separate from intensification, even though the two domains are obviously closely related by being two opposite poles of the wider function of degree modification or reference modulation.

While approximation is relatively understudied in morphology, it has been at the center of interest of other subfields of linguistics, such as pragmatics and discourse studies, which produced a huge amount of literature on the topic. In these fields, other terms are more common, some of which identify specific types of approximation. The most used terms are ‘mitigation’ (Caffi 2007), ‘indeterminacy’ (Bazzanella 2011), ‘imprecision’ (Balaș et al. eds. 2017), ‘hedging’ (Lakoff 1972; Kaltenböck, Mihatsch & Schneider 2010), ‘vagueness’ (Channell 1994; Mihatsch 2007), ‘intentional vagueness’ (Voghera 2012; Voghera & Collu 2017), or ‘defectiveness’ (Amiot & Stosic 2022). These studies, however, have not been paying much attention to morphological means to convey approximation (compared to other strategies such as discourse markers and particles), with the exception of diminutive markers used as attenuation strategies or being derived from approximative values (cf. Merlini Barbaresi & Dressler 1994; Merlini Barbaresi 2015; Grandi 2017). Therefore, there seems to be an overlapping zone between studies on approximation in general and evaluative morphology (see Fig. 1) that is in need of further and deeper investigation.
1.3 Research questions for this Special Issue

This Special Issue intends to bridge the gap between approaches to approximation in other subfields in linguistics and evaluative morphology, by concentrating on the expression of approximation by means of dedicated morphological means. The papers in this Special Issue are empirical studies of approximation in word formation from various theoretical perspectives, addressing one or more of the following research questions, each of which will be briefly discussed and illustrated in the following sections:

1. Which morphological means or forms are used to express approximating values?
2. What are the sources of approximating morphological markers crosslinguistically?
3. Which approximating values are expressed by morphological means?
4. Do we find competition between approximating morphological markers in a single language?
5. Do we find crosslinguistic tendencies/similarities in the morphological marking of approximation?

2. Forms

2.1 Morphological expression

As regards the morphological expression of approximation, suffixation (2) and prefixation (3) seem to be the predominant processes in the case studies in this Special Issue.
More marginally, **circumfixation** may also be used to express approximation, as can be seen in example (4) from Georgian.

(4) Georgian mžave ‘sour’ > mo-mžav-o ‘slightly sour’ (Topadze Gäumann 2015: 221–222)

In addition to affixes, **compound stems** or **affixoids** may also be used to express approximative values, as shown in the examples under (5).

(5) a. Italian simil- ‘similar’: simil-coppia ‘pseudo-couple’ (Masini & Micheli 2020)

b. French -forme ‘form’: une substance géliforme ‘a gel-like substance’ (Stosic & Amiot 2023)

c. Dutch imitate ‘imitation’: imitatieleer ‘imitation leather’ (Van Goethem & Norde 2020)

Finally, **reduplication** can be a morphological means to express approximation, as illustrated by examples (6a–b) (quoted in Masini & Di Donato 2023). Apart from contiguous reduplication, we find **discontinuous reduplication**, as in (6c) (Masini & Di Donato 2023).

(6) a. Makasar le’leng ‘black’ > le’leng-le’leng ‘blackish’ (Jukes 2006: 105)

b. Ma’di īngwē ‘white’ > īngwē īngwē ‘whiteish [sic], rather white’ (Blackings & Fabb 2003: 105)
c. Italian sapone-non-sapone lit. soap-NEG-soap ‘soap-free cleanser’ 
(Masini & Di Donato 2023)

2.2 Parts of speech as input and output categories

With respect to the bases (lexical categories) that serve as input of the morphological processes mentioned in the preceding section, the examples given thus far already show that nouns and adjectives are the most common bases. More examples are given in (7) for nouns and (8) for adjectives. However, approximative morphology can also be found in verbs (9), albeit to a lesser extent.

(7) a. Italian attore ‘actor’ > attore-non-attore actor-NEG-actor ‘unconventional actor’ 
(Masini & Di Donato 2023)

b. French frère ‘brother’ > frérot ‘bro’ (Stosic & Amiot 2023)

(8) a. English quasi-diplomatic (Cappelle, Daugs & Hartmann 2023)

b. Italian semi-aperto ‘partially open’ (Micheli 2023)

(9) a. Greek ψευτοκοιμάμαι (pseftokimáme) ‘I pseudo-sleep’ (Vassiliadou et al. 2023)

b. Italian quasi-cercare ‘sort of look for’ (Masini & Micheli 2020)

More marginally still, approximative morphemes may apply to adverbs (10a) and prepositions (10b):

(10) a. Sicilian cchiossai ‘more’ > cchjessanieddu ‘a little more’ (Brucale & Mocciaro 2023)

b. Italian simil-contro ‘sort of against’ (Masini & Micheli 2020)

It is worth noting that many strategies allow more than one base, like for instance English near, which attaches to both nouns (near-perfection) and adjectives (near-fatal; Cappelle et al. 2023). English -ish (Eitelmann & Haumann 2023) is probably the most flexible approximative morpheme in this respect. In addition to all previous base types, it may attach to full phrases (11a) or even completely debond (11b) (cf., among others, Norde 2009; Van Goethem & Norde 2020). This debonded use is however not exclusive for -ish. As shown in (11c), the neoclassical combining form pseudo-, for instance, also undergoes debonding in different languages, such as Swedish.

(11) a. English four years ago-**ish** (Eitelmann & Haumann 2023)

b. English I can cook. **Ish**. (Eitelmann & Haumann 2023)
c. Swedish [...] urbaniteten i fråga [...] blev först pseudo, sen sjangserade, och sen efterhärmandes ‘The urbanity under discussion [...] first became pseudo, then faded and then it was being imitated’ (Van Goethem, Norde & Masini 2021)

As regards the output categories of the morphological processes, homocategoriality between input and output category seems to be prevalent (in line with evaluative morphology properties; Grandi & Körtvelyessy 2015), especially with approximating prefixes as left constituents (12).

(12) a. French enquête ‘investigation’ > pseudo-enquête ‘pseudo-investigation’ (N > N) (Vassiliadou et al. 2023)
   b. Italian scientifico ‘scientific’ > parascientifico ‘parascientific’ (A > A); scheeggiarsi ‘chip’ > semi-scheeggiarsi ‘nearly chip’ (V > V) (Micheli 2023)

According to Stosic & Amiot (2023), homocategoriality is even defined as a constraint for diminutives to express approximation, see the opposition between (13a) and (13b):

   b. French bleu ‘blue’ > bleuet ‘cornflower’ (A > N)

However, word-class change in approximative word-formation is far from excluded, as shown in the contributions by Hüning & Schlücker (2023) on denominal adjectives in Germanic (14a) and by Treis (2023) on the derivational suffix -lab in Kambaata, which converts nouns, verbs and adjectives into adjectives (14b).

(14) a. English peacock > peacock-like feathers (N > A)
   b. Kambaata torr- ‘throw’ > torr-lab-á ‘in a kind of throwing way’ (V > A)

The English -ish morpheme again stands out here since it can take scope over a noun phrase (15) (Eitelmann & Haumann 2023).

(15) GEIST: How long have you been together?
     KOTB: [[A year and a half]DP-ish]DP. [COCA]

3. Sources

The approximating strategies illustrated in the previous section happen to derive from different sources that will be discussed in this section. What follows is a representative –
although not necessarily exhaustive – list of possible sources, most of which are explored in the articles hosted in this Special Issue, whereas others still need further exploration.

- **Negation items**: negative prefixes like *non-* in French (and possibly other languages) may convey approximation (e.g., *non-ville* ‘non-city’, namely a city with non-stereotypical properties, cf. Dugas 2017), but negative markers may also have a role in larger structures, like the Italian *N-non*-N discontinuous reduplication analysed by Masini & Di Donato (2023) in this volume (e.g., *sapone-non-sapone*, lit. soap-NEG-soap, ‘soap-free cleanser’).

- **Fake items**: items that convey meanings such as fakeness, imitation, simulation/pretending may easily turn into approximators. A representative example is the already mentioned *pseudo-*, from Ancient Greek *pseudēs* ‘false’, which is further analysed in this volume by Cappelle, Daugs & Hartmann (2023) and Vassiliadou et al. (2023). A possible example for the ‘pretending’ type is Finnish *-vinA* (*maksa-vina-an*, lit. pay-VINA-POS.3 ‘(they) pretend to pay’) (Salminen 2000).

- **Degree and quantity items**: this class can be split into at least two subclasses, namely: (i) items conveying an ‘almost’ or incompleteness meaning, e.g., *quasi-* in various European languages (e.g., English *quasiparticle*) (Cappelle, Daugs & Hartmann 2023); (ii) items meaning ‘half’, like Italian *semi-* (*semi-relazione* ‘pseudo-relationship’; Micheli 2023) or *mezzo-* (*mezzo-pacifiste* ‘half-pacifists’; Masini & Micheli 2020; Benigni 2022).

- **Spatial (proximity) items**: items that convey closeness may also become approximators, see for instance English *near-* (*near-synonyms, near-identical*; cf. Cappelle, Daugs & Hartmann 2023) or the neoclassical form *para-* (from Greek *para* ‘beside’), studied for Italian by Micheli (2023).

- **Diminutives**: the relation between diminutives and approximation is well-known (Jurafsky 1996); indeed, diminutives (unlike augmentatives, cf. Stosic & Amiot 2023) are commonly used to attenuate the meaning of the base (Dressler & Merlini Barbaresi 1994; Merlini Barbaresi 2015; Grandi 2017; also Voghera 2023), with different nuances depending on the lexical category and semantics of the base, cf. English *-let* in *star* (N) > *starlet* ‘a young actor likely to develop into a star’ (Dixon 2014: 172); French *-ette* in *réforme* ‘reform’ (N) > *réformette* ‘small reform with no value or scope’ (Amiot & Stosic
Similative items: elements meaning ‘like/as’, ‘seem/resemble’ or ‘similar/alike’ are commonly used as sources for approximation in the world’s languages (cf. Masini, Micheli & Huang 2018 for an overview); this is unsurprising given that “the semantic transition from similative comparison markers to approximation is extremely easy” (Mihatsch 2009: 79) since “similative comparison is never perfect, but always approximative” (Mihatsch 2009: 70, quoting Haspelmath & Buchholz 1998: 278). German -ähnlich ‘-like’ (ähnlich is an adjective meaning ‘similar’; Hüning & Schluesser 2023) and English -like (baptismal-like, prefix-like) are good examples. Bauer, Lieber & Plag (2013: 311–313) mention -like among similative adjective-deriving affixes (together with -esque, -ish and -oid) that have an approximating reading derived by inference from the similative one. Other examples come from Kambaata (Cushitic), e.g., -lab, from a verb root lab- ‘resemble, seem’ (e.g., gamball-(i)-lab-å(ta) ‘blackish’) (Treis 2023), and Georgian (Kartvelian), e.g., -savit, which includes the postposition vit ‘like/as’ (e.g., davighale-savit PRV-get_tired.AOR.1SUB-ES ‘I got like tired’; Topadze Gäumann 2015: 223).

Relational items: similarly to similative items, relational items – which express association and therefore, to some extent, similarity and belonging – may turn into approximators as well: this is the history Eitelmann & Haumann (2023) reconstruct for English -ish, which develops its current approximative meaning (cleanish, 50-ish) from a more relational/associative one (Spanish, heavenly).

Taxonomic items: these are possibly the best-known sources for approximation, giving rise to well-known markers of the sort/kind of type in a variety of languages (e.g., among many others, Tabor 1994; Denison 2002; Mihatsch 2007; Traugott 2008; Voghera 2013, 2017; Masini 2016). Taxonomic items also have a role in the evolution of morphological affixes, as is the case for German -artig (e.g., vogelartiges Tier ‘bird-like animal’), whose literal meaning is ‘belonging to a species (‘Art’)’ (Hüning & Schluesser 2023).
Modal items: finally, it is worth mentioning a source domain which is still underinvestigated, namely modality. In English and Italian, for instance, some epistemic adverbs are developing an approximative meaning in compound-like expressions: see, e.g., English maybe in a maybe-incident or Italian forse ‘maybe’ in forse-fidanzato ‘maybe-boyfriend’. The same is true of volitional expressions, like English wannabe (from want to be) in wannabe popstar, German möchtesegern ‘would very much like to’ in Möchtesegern-Schriftsteller ‘wannabe author’ or Italian vorrei-ma-non-posso ‘(I) would like but (I) can’t’ in boutique vorrei-ma-non-posso ‘wannabe boutique’ (cf. Norde et al. 2023). The conceptual step here is from an assessment of speakers’ (un)certainty about what they say (epistemic modality) or an aspiration to attain a specific property/state (volitional modality) to approximation.

4. Values

As may already be evident from the many examples given in the preceding sections, approximation is not a clear-cut domain and we would not expect anything different. As Ullmann (1962: 118) writes about the closely related notion of ‘vagueness’: “If one looks more closely at this vagueness one soon discovers that the term is itself rather vague and ambiguous”. The same can be said of approximation, which appears to be a complex functional domain comprising closely related yet distinct values, spanning from semantics to pragmatics.

In this section we identify and discuss some of these values (without any claim to exhaustiveness), focusing on those that emerge from the papers in the Special Issue, with special attention to approximating values that affect the propositional content (namely, what Voghera & Borges 2017 call “informational vagueness” vs. “relational/discourse vagueness”). As we will see, what emerges is that specific values may sometimes be quite difficult to define and that many markers may convey more than one value, which may generate quite some confusion.
The first notion that we consider is **privativity**, which Cappelle, Daugs & Hartmann (2023) define as a function that removes an essential property from an entity X, with the result that the output Y does not belong to X’s category anymore (‘not X’). For instance, *fake* in *fake blood* would have a privative function since *fake blood* does not match the properties of real blood, and therefore it falls out of the blood category. The privative function can therefore be regarded as the most extreme value of approximation, which implies the creation of a new (X-related) category rather than the modulation or evaluation of an existing category. See also the diachronic approximation cline proposed by Eitelmann & Haumann (2023), running from relational to privative (which is the meaning English -ish does not seem to realise).

Another value to be considered is **disproximation**, which is elaborated by Cappelle, Daugs & Hartmann (2023) in order to better characterize the (different) semantic contribution carried by the prefixes near-, pseudo- and quasi- in English. The authors advocate for a ‘dynamic’ or ‘orientational’ approach to meaning, which is implemented through Langackerian representations: whereas quasi- and especially near- are ‘approximating’, namely coming close to a standard (coinciding with the base X), pseudo- is ‘disproximating’ in that it falls short of something with respect to the standard, conveying an idea of deficiency which is not prominent in quasi-/near-.

In their semantic analysis of pseudo- in French and Greek, Vassiliadou et al. (2023) acknowledge its privative/disproximating value (category exclusion), but they also define two other non-privative values that pseudo- may convey, thus producing a three-way classification: the first coincides with a **subjective depreciation**, which implies category inclusion (see French pseudo-supporteur ‘pseudo-supporter’, denoting a team supporter who is a bad exemplar of the category for some reason and therefore judged negatively by the speaker); the second is vagueness intended as **undecidable categorization**, where category membership cannot be ascertained (see French pseudo-liberté ‘pseudo-freedom’). Category inclusion or exclusion is indeed a recurrent criterion throughout the papers in the Special Issue. See also Brucale and Mocciaro’s (2023) proposal to distinguish between:

---

1 Note that, in the typological literature, the term ‘privative’ normally refers to lexical and grammatical markers of absence like without or -less (cf. Stolz, Stroh & Urdze 2007), another equivalent term being ‘caritive’ (Oskolskaya et al. 2020).
(i) **internal approximation**, which entails a semantic modification of the base without altering category membership (see Sicilian *chilu* ‘kilo’ > *chiliceddu* ‘about one kilo’; *muzzucuni* ‘bite’ > *muzzucunieddu* ‘small bite’); (ii) **external approximation**, where the deviation from the standard is such that category membership is affected (see Sicilian *acqua* ‘water’ > *acquicedda* ‘sort of water’, but also more lexicalized cases like *figura* ‘figure’ > *figuredda* ‘holy card’).

Next, **attenuation** (or **reduction**) is probably one of the most widespread and quoted functions in the approximation domain. It is intended as the reduced degree of a quality or property and is especially common with color terms (French *verdâtre* ‘greenish’, English *whitish*, Italian *marrone* ‘brownish’, etc.) and other gradable adjectives (see (16a, c)). However, this function also applies to nouns (16b). In addition to reduced degree, attenuation may have to do with partialness or incompleteness, as the translations of the Italian examples (16a–b) illustrate.

(16) a. Italian *addormentato* ‘asleep’ > **semi-addormentato** ‘partially asleep’ (Micheli 2023)
   b. Italian *oscurità* ‘darkness’ > **semi-oscurità** ‘partial darkness’ (Micheli 2023)
   c. Kambaata *mux-i-láb-a-a<n>ta*
      *wet-EP-APRX-F.PRED-F.COP2<EMP>*
      ‘a bit wet, almost dry’ (Treis 2023)

Other values that are worth discussing are **fakeness, imitation** and **simulation**. As mentioned in Section 3, fake items have a role in the development of approximators, but **non-authenticity** can also be regarded as a specific value of approximation. The already mentioned *fake blood* is not a type of blood but still resembles blood because it reproduces some of its properties. On the other hand, as also pointed out by Cappelle, Denis & Keller (2018) (to which we refer for a finer-grained typology of ‘fakeness’), a *fake article* is still a type of article, despite deviating from the normal article in that its content is not true (as it should be). The question arises whether we are dealing with one single value (say, ‘non-authenticity’) or a family of related values. In their study of Italian *simil-*, Masini and Micheli (2020: 383) distinguish two different values: (i) fakeness, which is realized when “Y refers to an entity that is meant to imitate X (without being a genuine X) in order to be taken as an X” (e.g., *similpelle* ‘imitation leather’); (ii) imitation, which is realized when “Y refers to an entity that is meant to imitate or reproduce X (without being a genuine X), the
difference with [fakeness] being that there is no intent for Y to be taken as an X” (e.g., simil-vita, lit. SIMIL-life, referring to the virtual life of a videogame character). So, the distinction lies in the criterion of intentionality, which may produce different effects in terms of connotation. Both fakeness and imitation, as intended by Masini and Micheli (2020), are closely related to what we may call the simulative (or ‘pretending’) function, which is exemplified by the Finnish -vina- construction discussed by Ebner (2022) and exemplified in (17):

(17)  Pitä-ä osa-ta teeskennel-lä että ol-isì jotenkin
must-3SG be_able-INF pretend-INF that be-COND somehow
“teke-vinä-än työ-tät.”
do-VINA-POSS.3SG work-PART

‘One has to be able to pretend as if one was “seemingly working” somehow [although in reality there is no work]’

Beside fakeness and imitation, Masini & Micheli (2020) identify three other functions expressed by Italian simil-, namely: resemblance, (intentional) vagueness and kin-categorization.

Kin-categorization is realized when Y results in a separate category (with some sort of link to X) that gets conventionalized to a certain extent (e.g., Italian simil-matrimonio, lit. SIMIL-marriage ‘civil union’). This value has clear connections to both privativity (Cappelle, Daugs & Hartmann 2023) and external approximation (Brucale and Mocciaro 2023).

Resemblance, or similarity, is a central value that is instantiated by a variety of means and can be regarded as more basic and neutral than other values, because – in a way – approximation always involves some (level/type of) similarity between X and Y. It is not surprising, then, that resemblance or similarity is called upon, as a value, in many contributions within the Special Issue. If we take resemblance to be the value that emerges when “Y refers to an entity that is merely similar to X without being an X” (Masini & Micheli 2020: 383; e.g., Italian simil-Amazon ‘SIMIL-Amazon’, referring to an account that resembles the Amazon account but is not Amazon, namely an ‘Amazon-style’ account),

---

2 Ebner’s (2022) study of Finnish offers further insight into the (extension of the notion) of approximation. Beside the -vina construction, Ebner (2022) mentions the -maisilla- construction, which may define still another dimension of approximation, namely approximation as a function of time. Indeed, the -maisilla- construction is normally described in terms of “proximative aspect” (Ylikoski 2003: 43). The relation between (imminent or proximative) aspect and approximation definitely deserves further thought and investigation.
then we are in the domain of privativity and/or external approximation, although – as we already noted – category membership is not always so easy to determine. In her analysis of Kambaata’s marker -lab, Treis (2023) further distinguishes between “similarity and near-identity in manner” (e.g., uucc- ‘beg’ > uucc-lab-á ‘in a kind of begging way’) and “similarity in quality or character” (e.g., weteechchu ‘weteechchu tree’ > wet-eechch-lab-á ‘weteechchu-like’). Sometimes, resemblance is very close to imitation and distinguishing between the two values can be quite difficult (cf. Masini & Micheli 2020 on Italian simili- and Micheli 2023 on Italian semi-).

Hüning & Schlücker (2023) add yet another value to the picture, namely comparison (which has obviously to do with similarity). An -achtig formation in Dutch like leerachtig ‘leather-like’ in een leerachtige substantie ‘a leathery/leather-like substance’ can be both approximative (the substance has leather-like properties but is not leather) and comparative (the substance has leather-like properties regardless of its relation to the category leather), and differentiating between the two is not always easy or possible. The authors therefore recommend to distinguish approximative markers in the strict sense (which are privative and typically left-headed) from comparative markers with an optional approximative reading.

Last but not least, one of the main values expressed is (intentional) vagueness (Voghera 2012, 2013; Voghera & Collu 2017). This term refers to vagueness intentionally expressed by the speaker, not vagueness that depends on systemic factors (like bald being systemically a “vague predicate” because it gives rise to borderline cases, unlike for instance prime, when referring to numbers). Intentional vagueness has to do with uncertain or fuzzy categorization, with reference to entities that are (possibly) peripheral members of a category and whose categorial status is unclear. We already discussed the value of undecidable categorization as conceptualized by Vassiliadou et al. (2023). This may indeed be one subtype of intentional vagueness, the other being non-prototypicality, which expresses the peripheral status of an entity with respect to prototypical exemplars, due to some deviating property or properties to be reconstructed by the interlocutor (cf. Masini & Di Donato 2023). Non-prototypicality may result in category inclusion or exclusion, depending on individual cases.
Interestingly, expressions conveying intentional vagueness developed also a focus function, as illustrated for Italian diminutives by Voghera (2023). Thus, in example (18), the vague diminutive *un attimino* ‘an instant DIM’ puts the following NP (‘this draft law that is being discussed’) in focus.

(18)  
\begin{quote}
vi volevo aggiornare un attimino su questo disegno di legge che è in discussione
\end{quote}

‘I wanted to update you a little instant on this draft law that is being discussed’

In the background of all these values lies a bigger question: What is the relationship between approximation and categorization? Where does approximation end and where does categorization start (or vice versa)? When we use expressions like *blueish* or *prefix-like* are we approximating an existing category or concept or rather creating a new one? In response to a question like *What’s your favourite shade of blue?*, we are likely to reply something like *baby blue* or *dark blue*, but not *blueish*. Still, *blueish* has a conceptual value, closely related to *blue*, although possibly peripheral with respect to the prototypical core of the category. Whereas the concept *baby blue*, as a specific shade of *blue*, is probably more stable and shared by a large number of people, *blueish* is more variable and subject to speakers’ interpretation. Moreover, it probably blurs into other closely related concepts like for instance *greenish*, due to the fuzzy boundaries of color categories (see Fig. 2).

![Fig. 2: Blueish vs. baby blue](image)

To sum up, we may try to visualize the – admittedly complex – picture described in this section as in Fig. 3. The figure schematizes the continuum going from category modulation or ‘evaluation’ (here, specifically, through approximation) to (new) category creation.
Fig. 3: From approximation to categorization

The boundaries between the three areas – red (indicating the standard/norm/prototype, which coincides with the base X), yellow (representing the deviations from the standard/norm/prototype) and grey (indicating the conceptual area beyond the category X) – are intended as fuzzy (hence the dotted white line), in line with prototype semantics: it is not always clear if an approximated item Y still belongs to the standard’s category X or not. This representation is in harmony with: (i) the view of evaluation (and hence also approximation) as a process that expresses a deviation from a standard or default value (Grandi 2017); (ii) the view that regards approximation and categorization as two sides of the same coin (cf. Vassiliadou et al. 2023; Vassiliadou & Lammert 2022). Finally, deviations from the standard (the yellow area) is purposely oversimplified in Fig. 3: as is clear from the above discussion, deviations may be of different types and intensity, generating a range of approximating values, also depending on the nature of the base X.\(^3\) Hence, more fine-grained schemas are needed for each value, some of which are provided by the papers in this Special Issue. Note that the standard/norm/prototype may be recursive: baby blue, for instance, is a hyponym of blue, which may well develop its own deviations (e.g., babyblueish, near-babyblue, etc.).

---

\(^3\) One direction that is not explored in the Special Issue but would be worth pursuing is the hypothesis that some sources (Section 3) may tendentially develop into approximation markers of some sort only, thus constraining the limits of possible diachronic development. According to Masini, Micheli & Huang (2018), approximation markers derived from simulative items tend to convey only a subset of the possible values: they can end up expressing intentional vagueness or imitation, whereas attenuation seems to be excluded or dispreferred (a claim to be tested on a wider sample of languages).
5. Competition

Approximative constructions often find themselves in competition with each other, at different levels.

First of all, we may find different structural strategies competing with each other, for instance X-wannabe (actress wannabe) versus wannabe-X (wannabe actress) (Norde et al. 2023).

Secondly, different approximative morphemes can be found with the same base, as is shown by examples (19a) and (19b), in which the German approximative suffixes -ähnlich and -artig both collocate with Honig ‘honey’ to form denominal adjectives (Hüning & Schlücker 2023).

(19) a. Eine honigähnliche Zuckerpaste wird gegen die natürliche Wuchsrichtung des Haares aufgetragen und dringt so bis zum Haarschaft ein.

‘A honey-like sugar paste is applied against the natural direction of hair growth, penetrating all the way to the hair shaft.’

b. Wer es süß mag, kann die frischen Blüten zu einem Gelee oder einem honigartigen Sirup für einen Brotaufstrich verarbeiten.

‘For those who like it sweet, the fresh flowers can be made into a jelly or a honey-like syrup for a spread.’

In other cases, competing morphemes occur in very similar collocations, for instance in the food items in (20) (Eitelmann & Haumann 2023).

(20) Is everything in here onion-like, bacon-esque or cheese-ish?

Nevertheless, competing approximative morphemes, even when they appear in similar contexts, may show distributional preferences (Hüning & Schlücker 2023) and/or semantic differences. For example, Bauer, Lieber & Plag (2013: 416) argue that the prefixes pseudo- and quasi- and the suffixes -ish, -esque and -oid all denote “something that is similar, but not identical, in shape or quality to what the base denotes”, although only pseudo- conveys a notion of ‘falseness’. As far as the suffixes are concerned, they claim that these may differ according to register or domain without further specification. But as Cappelle, Haugs & Hartmann (2023) show, it is possible to reveal differences using methods from distributional semantics, such as semantic vector spaces (for another semantic vector
analysis, of Dutch compounding elements meaning ‘fake’, see Van Goethem & Norde 2020).

In the paper by Treis in this Special Issue, on the other hand, the (near-)synonymy of approximative morphemes is being disputed. Treis (2023) quotes a Kambaata dictionary entry (Alemu 2016: 340) containing a compounding element *manka’-á* ‘humility, behavior, manner(s), norm(s)’, whose definition contains the approximative suffix *-lab*, as illustrated in (21). However, most of the dictionary entries with *manka’-á* were unknown to Treis’ language assistant, nor did they appear in elicited data, so either there is no competition, or there is but is limited to specific varities of Kambaata.

\[
(21) \quad \begin{array}{l}
\text{baar}-\text{manka’-á} \\
\text{yellow-manner-M.ACC} \\
\text{feeg-á} \\
\text{bright-M.ACC} \\
\text{baar-á} \\
\text{yellow-M.ACC} \\
\text{ih-umb-á;} \\
\text{become-3M.NEG5-M.ACC} \\
\text{baar-lab-á} \\
\text{yellow-APRX-M.ACC}
\end{array}
\]

‘yellowish’ [Definition:] ‘not being bright yellow, yellowish’

A third level of competition is one between different variants of the same morpheme, such as Modern Greek *pseudo*- and *psefto*- (Vassiliadou et al. 2023) in ψευτο/ψευδο-επιστήμη (psefto/psevðo-epistími) ‘pseudo-science’ or ψευτο/ψευδοδίλημμα (psefto/psevðoðílima) ‘pseudodilemma’. As Vassiliadou et al. (2023) show in their analysis of Modern Greek corpus data, *pseudo*- and *psefto*- do however show distributional differences, e.g., in the parts of speech they collocate with (with more variation for *psefto*-).

Finally, approximative morphemes may compete with other (non-morphological) constructions, e.g., expressions such as English *kind of*/*kinda* and *sort of*/*sorta* (Eitelmann & Haumann 2023), or the Sicilian degree adverb *tanticchia/nna picca* ‘a bit’ (Brucale & Mocciaro 2023).

Factors determining these various kinds of competition are largely a topic for future research, but early studies indicate that strongly lexicalized approximative items may block the use of other approximators, as in the case of Dutch *kunstmest* ‘fake manure > fertilizer’, which blocks the use of words like *imitatiemest* (‘imitation manure’). What is most probably at play here is what Goldberg (2016: 377–378) has termed “statistical pre-emption”, which she defines as “a particular type of indirect negative evidence that results from repeatedly hearing a formulation, B, in a context where one might have expected to
hear a semantically and pragmatically related alternative formulation, A.” The choice between competing approximative morphemes can also be constrained by semantic factors. Thus, the sense of pseudo- of ‘pretending to be without actually being’ precludes collocations such as 'pseudo-fatal' (Cappelle, Haugs & Hartmann 2023). Yet other factors remain to be explored. These include the etymology of the base (e.g., neo-classical compounds vs native bases) or register.

6. Cross-linguistic observations

The studies in this Special Issue are drawn from a fairly large (albeit unbalanced) sample of languages: Germanic (English, Dutch, German), Romance (Italian, French, Sicilian), Greek, Serbian – all IE languages – and Kambaata (Cushitic). From this sample, some very general cross-linguistic tendencies can be derived:

i) morphological means to express approximation seem to be widely spread across European languages and beyond;

ii) there is a tendency for approximative affixes to accommodate different kinds of bases, especially nouns, adjectives and verbs, but other categories are not excluded;

iii) the types of sources are also quite recurrent among languages, with single languages often displaying multiple markers from different sources;

iv) languages often display an array of approximating values, sometimes conveyed by the same marker(s), sometimes expressed through dedicated strategies; disentangling the exact relationship between forms and functions and how they impact the choice between competing expressions seems to be one of the most challenging tasks that lie ahead.

The similarity displayed by European languages in the morphological expression of approximation is also due to their common heritage from classical languages (see neoclassical combining forms such as pseudo-, quasi-, para-, semi-, etc.). Nevertheless, as shown in the cross-linguistic study by Van Goethem, Norde & Masini (2021) on the prefix pseudo-, the same approximative morpheme may have different fate and behavior cross-linguistically. Although that study indicates that pseudo- displays a similar lexical distribution in Romance and Germanic languages, notable differences in productivity, for
instance, are detected. Fig. 4 shows the “global productivity” (Baayen and Lieber 1991: 818–819) of pseudo- in the eight languages under examination, with potential productivity (i.e. the number of hapaxes divided by the number of tokens) on the y-axis and type frequency on the x-axis. The plot distinguishes between the Romance and the Germanic languages (the former written in red characters, the latter in grey), and between nominal and adjectival bases (the former red-dotted, the latter blue-dotted). It can be derived from the graph that the differences in productivity are not only related to the language group (overall higher productivity in Romance than in Germanic) but also to the base types (overall higher productivity in nominal than in adjectival formations).

Fig. 4: Global productivity of pseudo-

To explain these kinds of differences, intralinguistic competition is one of the factors to be taken into account (cf. also Cappelle, Daugs & Hartmann 2023). Crucially, there is not always a one-to-one mapping between approximative morphemes in different languages. The contribution by Hüning & Schlücker (2023) nicely illustrates this point for Germanic denominal suffixation: it shows that three different German approximative suffixes correspond to Dutch -achtig (and English -like), namely -artig, -ähnlich and -haft. These compete with each other and show a similar high frequency of use, albeit with important lexical-distributional differences. More fine-grained contrastive analyses of this kind are
needed to further explore the universal and divergent tendencies in the field of approximative morphology from a cross-linguistic perspective.

References


Approximation in Morphology: A State of the Art

Francesca Masini
Alma Mater Studiorum – Università di Bologna
Dipartimento di Lingue, Letterature e Culture Moderne
Via Cartoleria 5
40124 Bologna, Italy
francesca.masini@unibo.it

Muriel Norde
Nordeuropa-Institut
Humboldt-Universität zu Berlin
Unter den Linden 8
10099 Berlin, Germany
muriel.norde@hu-berlin.de

Kristel Van Goethem
F.R.S.-FNRS & Université catholique de Louvain
Institut Langage et Communication / PLIN / VALIBEL
Place Blaise Pascal 1
1348 Louvain-la-Neuve, Belgium
kristel.vangoethem@uclouvain.be

This is an open access publication. This work is licensed under a Creative Commons Attribution CC-BY 4.0 license. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/