HIGH AND LOWS OF THE FRENCH DIMINUTIVE SUFFIX *-ET(TE)*[1]

ADÈLE HÉNOT-MORTIER

ABSTRACT: French assigns grammatical gender (masculine or feminine) to nominals, and is endowed with a diminutive suffix *-et*/*-ette*. In most cases, the diminutive noun resulting from *-et(te)*-affixation will have the same gender as its base (Bally 1932), but there are a significant number of exceptions to this rule, that most of the previous literature (Dauzat 1937; Milner 1989 i.a.) took to be the result of lexicalization. In this study, we assess how frequent gender mismatches induced by *et(te)*-affixation are, in either direction (masculine to feminine and vice-versa), and what the exact semantic consequences turn out to be. In particular, we show that there exists a significant frequency asymmetry between *-et*-affixation and *-ette*-affixation, which affects both gender-matching and gender-mismatching base-derivative pairs, supporting the idea that gender-mismatching diminutives are to a certain extent morphologically transparent, but also that *-ette*-affixation may receive an analysis distinct from that of *-et*-affixation. We provide a analysis within the framework of Distributed Morphology (Halle & Marantz 1993) that is in line with the statistical data and with recent cross-linguistic findings on diminitive an augmentative affixes, according to which such elements may vary in place and manner of attachment, across, and also within, languages (Wiltschko & Steriopolo 2007).

KEYWORDS: diminutives, morpho-semantics, statistical analysis, word embeddings, distributional semantics

1. BACKGROUND ON DIMINUTIVE AND AUGMENTATIVE AFFIXES

1.1 Derivation or inflection?[2]

Diminutive and augmentative affixes (henceforth D&A affixes) tend to express a difference of size with respect to the base they attach to, and/or a relation of endearment/contempt (Stump 1993, Jurafsky 1996 i.a.). Such affixes have been

---

[2] Many thanks to an anonymous reviewers for helping me work out the next couple of sections.

a puzzle for morphological theory partly due to the difficulty to unequivocally classify them into derivational vs. inflectional morphology (Scalise 1986, Stump 1993 i.a.). More specifically, the fact that such affixes can be iterated (cf. (1)) and create new meanings advocates for a derivational analysis, while the fact they are to some extent transparent with respect to syntactic features (such as gender, cf. (2)) and/or categories (cf. (3)) advocates for an inflectional analysis.

(1) Repeated application (Zulu, Doke 1930, taken from Stump 1993)

intombi → *intomb-azana* → *intomb-azany-ana*
'maiden' → 'little girl' → 'very small girl'

(2) Gender preservation (Breton, Stump 1993). F-indices indicate the noun is grammatically feminine, M-indices, that it is masculine.

a. *potr$_M$* → *potr-ig$_M$*   b. *merc'h$_F$* → *merc'h-ig$_F$*
   'boy' → 'little boy'   'girl' → 'little girl'

(3) Category preservation (Russian, Wiltschko & Steriopolo 2007)

a. *xap-á-t'* *→ *xap-ok-e*
   'to grab' *→ 'person who grabs a little'

b. *žád-n-ij* *→ *žad-ok-e*
   'stingy' *→ 'a person who is a little stingy'

Scalise (1986) proposed that evaluative affixes (understood as feature-preserving affixes), were forming a separate layer of morphological derivation, in-between derivation and inflection. This view turned out to be unsatisfactory for three reasons: (i) it does not account for all kinds of uses of D&A affixes, in particular cases in which they are not necessarily transparent with respect to syntactic features or even categories; (ii) it does not explain why in the first place evaluative affixes have this special status; and, (iii) it does not fully reflect the cross-linguistic picture, which reveals that evaluative affixes may not be sandwiched between derivational and inflectional morphology,[3] and that, moreover, other kinds of morphological elements (e.g. compound verbs) pattern like evaluative affixes. Based of these observations, Stump (1993) proposed an amendment of Scalise's idea to deal with problem (ii) and (iii), based on the assumption that all evaluative affixes have the core property of being category-preserving (without necessarily preserving *all* syntactic features, or being of the D&A kind). However, this analysis still had to stipulate the

---

[3] Stump reports examples from Southern Barasano, Welsh, Kikuyu showing that diminutives can follow plural inflections. Similar data can be found in Kolyma Yukaghir and Itelmen (Steriopolo 2017).

existence of three classes of morphological transformations (grouping together rules depending on their input-output signature), one of them being the class of category-preserving transformations. Additionally, the analysis left point (i), along with its consequences for the semantics attached to D&A affixes, unaccounted for.

1.2 The view from Distributed Morphology

The emergence of Distributed Morphology (henceforth DM, Halle & Marantz 1993, 1994) offered a new framework to capture the puzzling behavior of D&A affixes in term of both morphology and semantics. As a brief reminder, the main assumption behind DM is that there is no division between syntax and morphology; the functions that other theories ascribe to the "lexicon" are "distributed" among other components of the grammar. In that framework, the so-called *Formative List* provides the input for syntax in the form of feature bundles (devoid of phonological features). Syntactic operations, such as MERGE, MOVE, and AGREE, apply to formatives. The so-called *Exponent List* contains the phonological forms compatible with each formative, and is used for (post-syntactic) vocabulary insertion. Importantly in DM, roots (which are taken to be formatives) are proposed to be category-neutral, and must undergo categorization by the action of *heads*: *n* (nominalizing head), *a* (adjectivizing head), *v* (verbalizing head) etc. This is exemplified in Figure 1.

(a) THE NOUN 'HAMMER'   (b) THE VERB 'TO HAMMER'

FIGURE 1: CATEGORIZING A ROOT IN DIFFERENT WAYS IN DM

Subsequent work within the DM framework (Marantz 1997; Harley & Noyer 1998; Marantz 2001; Arad 2005, 2005), points out the crucial distinction between creating words from roots and creating words from existing words, that is, from roots that are already merged with some word-creating head. This is known as the Lexical Decomposition Hypothesis.[4] Behind this division is the idea that merging a head creates a new "opaque" morphological object whose internal structure and properties cannot be accessed by subsequent morphological operations. In particular, the first head to be merged takes a

---

[4] This hypothesis is often associated with DM, but not restricted to this framework (Borer 2005, 2013; Fathi & Lowenstamm 2016).

morpho-phonologically and semantically underspecified root and assigns it a form and a meaning that will be the only things available to further transformations; in other words, the first head is supposed to "set" the semantic and morpho-phonological features of the newly created word. This can successfully explain the existence of opacity effects witnessed in *both* the semantic and the phonological domain. For instance, affixes such as *-ity*, which shift the stress assignment pattern of the base they attach to, also tend to lead to semantically less predictable meanings than otherwise similar affixes, such as *-ness*, which do not shift stress (Aronoff 1976). This can be captured by assuming that the former kind of affix merges at the root-level (where several possible forms and meanings are still accessible and modifiable), while the latter kind merges with already created words, whose phonology and semantics are more fixed. This is illustrated in Figure 2.
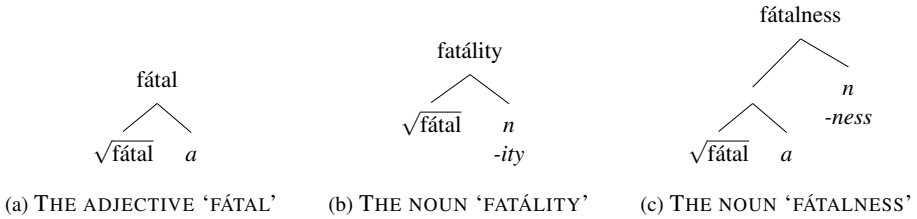

(a) THE ADJECTIVE 'FÁTAL'   (b) THE NOUN 'FATÁLITY'   (c) THE NOUN 'FÁTALNESS'

FIGURE 2: DERIVING *fatálity* (ROOT-LEVEL) AND *fátalness* (ADJECTIVE-LEVEL) FROM THE ROOT √FÁTAL

Why is this interesting for the analysis of D&A affixes? By assuming those affixes are either modifiers or nominalizing heads that may merge at the root-level or at the word-level, Wiltschko & Steriopolo (2007) derived four different clusters of properties for D&A affixes.[6] The root-level vs. word-level parameter controls whether or not an affix can seemingly apply to various syntactic categories (in fact, the same uncategorized root) and whether or not it can appear above other derivational markers – in particular, root-level D&A affixes. The head vs. modifier parameter controls whether or not an affix can modify the morpho-syntactic features (category, gender,[7] mass/count, inflectional class etc.) of its input. This is summarized in Table 1. Interestingly, neither the root

---

[6] Other accounts, e.g. (De Belder *et al.* 2014), simply assumed that D&A affixes could be below or above the level of categorization, without a head/modifier distinction. However, it is hard to see how these accounts capture *at the same time* (i) cases in which affixes impose gender but appear category-specific (signaling word-level derivation), like *-chen* in German (Wiltschko & Steriopolo 2007); and (ii) cases in which affixes do *not* impose their features but merge above number (signaling again word-level derivation), like *-die* in Kolyma Yukaghir or *-čχ* in Itelmen (Steriopolo 2017). This calls for another parameter controlling feature transparency.

[7] As we will focus on the interaction between diminutization and gender in the next Sections,

vs. word nor the head vs. modifier parameter appears to be a language-level parameter, since languages like Russian exhibit D&A suffixes varying across both dimensions (Steriopolo 2017). Italian, Modern Hebrew, Polish, Spanish, and Tunisian Arabic were also argued to exhibit language-internal variation of the root vs. word parameter (De Belder *et al.* 2014), some of these languages (e.g. Italian) even using the *same* surface suffix to realize both variants.

	Root-level head	Root-level modifier	Word-level head	Word-level modifier
Applies to diverse categories	✓	✓	✗	✗
Preserves input morpho-syntactic features	✗	✓	✗	✓
Can appear above inflection	✗	✗	✗	✓
Can appear above derivation	✗	✗	✓	✓
Examples	Russian *-ug/-an*	Halkomelem RED	German *-chen/-lein*	Kolyma Yukaghir *-die*; Itelmen *-čχ*; Russian *-k/-c*

TABLE 1: FOUR CLASSES OF D&A AFFIXES PREDICTED BY THE DM FRAMEWORK OF WILTSCHKO & STERIOPOLO (2007); STERIOPOLO (2017); STERIOPOLO *et al.* (2023).

To the above predictions we would like to add the following ones, summarized in Table 2: root-level categorization (i.e., merger of a *head*) should be less productive and semantically compositional than word-level derivations in general (a point already made for D&A affixes by De Belder *et al.* 2014, and for verb formation in Modern Hebrew by Arad 2003), but also, we think, than root-level *modification*, due to the fact that modification is not constrained by selection, nor creates opacity.[8] More specifically, if $\sqrt{r}$ is a root, A a D&A affix in the form of a root-level head, and if the nominalization of $\sqrt{r}$ has meaning M then, $A(\sqrt{r})$ should not necessarily mean *small/big M* or even express endear-

---

we want to make clear that the background assumption here is that gender features are neither represented in the category-neutral roots, nor in a dedicated projection, but instead, are hosted in nominalizing heads (Lecarme 2002; Ferrari 2005; Kihm 2005; Acquaviva 2009; Kramer 2015, 2016; Deal 2016). This includes diminutive and augmentative heads.

[8] If we assume that, from a semantic perspective, modification just intersects the sets of possible referents of the root with the property P denoted by the modifier, then root-level modification just returns all the possible referents of the root *that are P*, i.e. returns a subset of the original root that retains some degree of underspecification. When this root gets subsequently categorized, the corresponding output is then semantically equivalent to that of the categorized *and then* modified, root. In other words, this predicts that categorization by a head and modification should commute at the semantic level.

ment/contempt towards instances of M. Instead, $A(\sqrt{r})$ is expected to express diminutization/augmentation towards something vaguely related to M.

	Root-level head	Root-level modifier	Word-level head	Word-level modifier
Productive	✗	✓	✓	✓
Semantically compositional	✗	✓	✓	✓

TABLE 2: TWO ADDITIONAL PREDICTIONS REGARDING THE FOUR CLASSES OF D&A AFFIXES.

Figure 3 summarizes the four possible structures and behaviors of D&A affixes, assuming the affix is diminutive (DIM), and that the output is nominal. DIM_h stands for a head-diminutive, while DIM_m stands for a modifier-diminutive. Nodes are associated with features and/or meanings. $\tilde{M}$ refers to the underspecified meaning of the root $\sqrt{r}$, that gets fixed to a more specific nominal meaning M *via* n-categorization. $\mathscr{F}$ stands for features; $\mathscr{F}_{\text{DIM}}$ are the features carried by the DIM-head, $\mathscr{F}_n$, by another n-head.
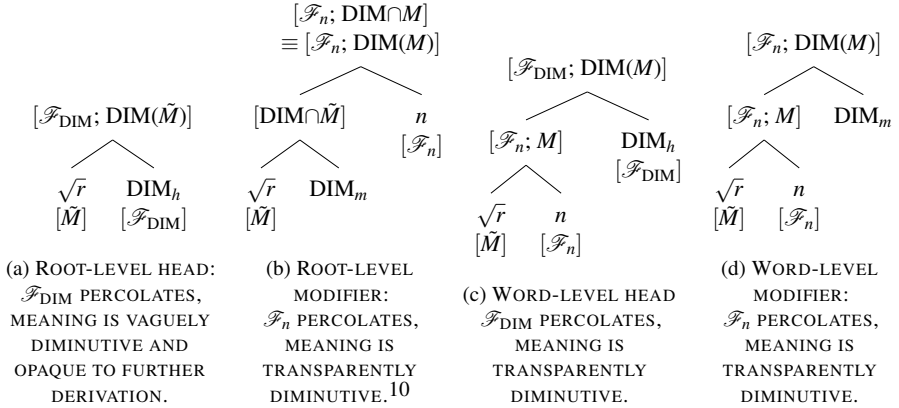

FIGURE 3: SUMMARY OF THE 4 POSSIBLE STRUCTURES ASSOCIATED WITH D&A AFFIXES, AND OF THEIR CONSEQUENCES ON THE FEATURES AND MEANING OF THE OUTPUT WORD.

The rest of this paper is structured as follows. In Section 2, we give an overview of the intricate and sometimes conflicting properties of the French diminutive *-et/-ette*, showing that these two putative variants of the same suffix follow distinct patterns and thus may, in some cases at least, instantiate distinct abstract morphological elements. We bring support for this observation *via*

---

[10] Footnote 8 explains why we think it is the case.

a more systematic analysis of the French lexicon in Section 3, focusing on gender-matching vs. gender-mismatching nominal diminutives. In particular, we find that -*ette*-suffixation is more frequent than -*et*-suffixation *across the board*, meaning, for both gender-matching base-derivative pairs, and gender-mismatching ones. We also show that the proportion of "true" diminutive derivatives is dependent on two parameters: whether or not the pair is gender-matching and whether or not the derivative is feminine. This leads us to recast the empirical observations in more definitive and formal terms in Section 4, where we argue that the frequency and meaning differences observed can be explained (as well as most of the observations from Section 2) if we assume that the pair -*et/-ette* is ambiguous between realizations of a transparent, root-level diminutive modifier, and realizations of two distinct heads with a slightly different semantics, one masculine and root-level, and the other, feminine and word-level.

2. THE FRENCH DIMINUTIVE SUFFIX -*ET/-ETTE*

Among other D&A affixes, French is endowed with a "diminutive" suffix -*et/-ette*. The exact nature of the diminutive semantics expressed by this suffix is notoriously hard to describe in a uniform fashion, as pointed out by e.g. Hasselrot (1957, 1972); Weber (1963); Milner (1989); Delhay (1995); Fradin (2003). We now proceed to analyze this suffix in light of the observations brought up in the previous section.

2.1 Feature preservation

French assigns grammatical gender (masculine or feminine) to all nominals. Feminine bases generally yield feminine -*ette*-diminutives, while masculine bases generally yield masculine -*et*-diminutives (Bally 1932). This is shown in (4).

(4)   a.   *maison$_F$* → *maisonn-ette$_F$*
            'house' → 'small (cute) house'

    b.   *balcon$_M$* → *balconn-et$_M$*
            'balcony' → 'small (cute) balcony'

However, Milner (1989) observed that -*ette* may attach to masculine bases and -*et* to feminine bases – a phenomenon we dub *gender-mismatch* – leading to a looser semantic relationship between the base and the derived form. The pairs in (5) and (6a) exemplify this observation. Moreover, it appears that certain

bases, such as *boule* ('ball'), can be suffixed with both *-et* and *-ette*, leading to a gender-matching, diminutive pair (6b), and a gender-mismatching, non-diminutive pair (6a). This optionality suggests that *-et/-ette* may be ambiguous between a semantically and featurally transparent modifier and a semantically and featurally opaque root-level head.

(5) $char_M \rightarrow char\text{-}ette_F$
$char_M \quad *\rightarrow char\text{-}et_M$
'chariot' $\rightarrow$ 'cart'

(6) a. $boule_F \rightarrow boul\text{-}et_M$
'ball' $\rightarrow$ 'cannonball'/'ball'(and chain)

   b. $boule_F \rightarrow boul\text{-}ette_F$
'ball' $\rightarrow$ 'small ball'

2.2 Productivity

The feminine variant of the suffix is relatively productive, while the masculine variant is sensibly less so, as shown by the examples in (7) based on recently adopted English loanwords.[12] The successful derivations are all gender-matching and associated with a meaning of the form *small/insignificant M* where *M* is the meaning of the non-diminutized noun. This constitutes evidence that *-ette* is somehow more likely than *-et* to be a productive morpheme (either a modifier or word-level head).

(7) a. $covid_{M/F}$[13] $?\rightarrow covid\text{-}et_M$
$\rightarrow covid\text{-}ette_F$

 l. $brunch_M ?\rightarrow brunch\text{-}et_M$

   b. $start\text{-}up_F \rightarrow start\text{-}up\text{-}ette_F$

   c. $deadline_F \rightarrow deadline\text{-}ette_F$

   d. $review_F \rightarrow review\text{-}ette_F$

   e. $punchline_F \rightarrow punchlin\text{-}ette_F$

   f. $milestone_F ?\rightarrow mileston\text{-}ette_F$

   g. $timeline_F ?\rightarrow timelin\text{-}ette_F$

   h. $vibe_F ?\rightarrow vib\text{-}ette_F$

   i. $hype_F ?\rightarrow hyp\text{-}ette_F$

   j. $update_F *\rightarrow updat\text{-}ette_F$

   k. $keynote_F *\rightarrow keynot\text{-}ette_F$

---

[12] Judgments were produced by the author.

m. *lunch*$_M$?→ *lunch-et*$_M$ t. *job*$_M$ *→ *job-et*$_M$

n. *workshop*$_M$?→ *workshop-et*$_M$ u. *burnout*$_M$ *→ *burnout-et*$_M$

o. *process*$_M$ *→ *process-et*$_M$ v. *challenge*$_M$ *→ *challeng-et*$_M$

p. *feedback*$_M$ *→ *feedback-et*$_M$ w. *buzz*$_M$ *→ *buzz-et*$_M$

q. *conf-call*$_M$ *→ *conf-call-et*$_M$ x. *week-end*$_M$ *→ *week-end-et*$_M$

r. *chart*$_M$ *→ *chart-et*$_M$ y. *benchmark*$_M$ *→ *benchmark-et*$_M$

s. *meeting*$_M$ *→ *meetingu-et*$_M$

2.3 Range and domain

As noted in e.g. (Dal 1997), the *-et/-ette* suffix surfaces on nominals, but also on adjectival bases (to create new adjectives, cf. (8)), and verbal bases (to create deverbal nouns denoting tools associated with the action denoted by the corresponding verb, cf. (9)). As (8f) vs. (9h) however illustrate with recent loanwords, adjectival derivations are not very productive while verb-to-noun derivations are.[14] This nevertheless shows that *-et/-ette* can apply to various syntactic categories (suggesting the availability of root-level derivation) but sometimes preserves and sometimes changes, this input category (suggesting optionality between modification and categorization). Moreover, in the verb-to-noun case, the preferred suffix appears to be the feminine one, which implies that either *-ette* is more likely to be a root-level head, or that feminine gender is a default in cases of modification, when no gender features are available.[15] In the adjective-to-adjective case, gender is meant to be determined *via* agreement with the noun the adjective combines with.

(8)   a. *mignon* → *mignon-et(te)*   d. *gentil* → *gentill-et(te)*
       'cute' → 'cutesy'   'nice' → 'niceish'

      b. *propre* → *propr-et(te)*   e. *jeune* → *jeun-et(te)*
       'clean' → 'cutely clean'   'young' → 'youngish'

      c. *long* → *longu-et(te)*   f. *cringe* *→ *cring-et(te)*
       'long' → 'longish'   'young' → 'youngish'

(9)   a. *balayer* → *(balay-ette)*$_F$ c. *trottiner* → *(trottin-ette)*$_F$
       'sweep' → 'brush'   'jog' → 'kick scooter'

      b. *pousser* → *(pouss-ette)*$_F$ d. *laver* → *(lav-ette)*$_F$
       'push' → 'stroller'   'wash' → 'dishcloth'

---

[14] Note the proper term for *vape* in French is *vapoteuse*; however, *vapotette* looks like another likely candidate, that, we think, could be understood as *vape* by most speakers.

[15] We will adopt this last hypothesis in our analysis.

e. *taper* → *(tape-ette)$_F$* g. *souffler* → *(souffl-et)$_M$*
 'hit' → 'swatter'   'blow' → 'bellow'

f. *siffler* → *(siffl-et)$_M$* h. *vapoter* → *(vapot-ette)$_F$*
 'whistle' → 'whistle'   'vape' → 'vape'

2.4 Interaction with other morphology

-ette may surface after some derivational morphology (10), but always before inflectional morphology such as pluralization (11a). It was hard to find examples of *-et* after derivation, but it is certain that *-et* too surfaces before inflection (11b).

(10) a. *révolu-tion$_F$* → *révolu-tion-ette$_F$*
   'revolution' → 'small/insignificant revolution'

  b. *promo-tion$_F$*   → *promo-tion-ette$_F$*
   'discount/promotion' → 'small/insignificant discount/promotion'

(11) a. *maison-s$_F$* *→ *maison-s-ette$_F$*
   'houses' → 'small (cute) houses'

  b. *balcons-s$_M$* *→ *balcon-s-et$_M$*
   'balconies' → 'small (cute) balconies'

Regarding combinations of more than one D&A affix in French, the possibilities are perhaps unexpectedly restricted.[17] First, it appears impossible to iterate *-et/-ette*, as shown by the first line of (12a) and (12b). It is yet possible to achieve a similar intensified diminutive meaning using *-inet/-inette*. If *-in/-ine* can probably be analyzed as a proper suffix in French (commonly found at the end of feminine proper names, and names of chemical compounds), it is unclear whether it is really a productive *diminutive* suffix; and, at the very least, it does not seem to be so in the cases where *-inet/-inette* is licensed, as shown by (13).

(12) a. *maison-ette$_F$* *→ *maison-et(te)-ette$_F$*
       ?→ *maison-inette$_F$*
   'small house' → 'very small house'

  b. *bomb-ette$_F$* *→ *bomb-et(te)-ette$_F$*
       → *bomb-inette$_F$*
   'small bomb' → 'very small bomb'

---

[17] I would like to thank an anonymous reviewer, as well as Marcel den Dikken, for raising this point, which turns out to be a weak spot of my analysis. I am aware that the blocking-based explanation I put forth in this paragraph is unsatisfactory, and will try to look better at other corners of French as well as at the cross-linguistic picture to hopefully better answer this puzzle in the future.

(13)   a.   *maison$_F$* *→ *maison-ine$_F$*   b.   *bombe$_F$* → *bomb-ine$_F$*
        'house'   *→ 'small house'   'bomb' → 'potato dish'

These data might explain the impossiblity of double *-et/-ette* diminutization in French. If *-inet/-inette* is analyzed as one single morpheme, whose meaning is that of an intensified diminutive, then its competition with iterated *-et/-ette* suffixation might be the explanation as to why the latter seems to be blocked in French.

2.5 Summary

Table 3 summarizes the somewhat intricate conclusions that can be drawn about *-et/-ette* from each of the previous subsections.

		-et	*-ette*	Diagnoses the possibility of...
(i)	Applies to diverse categories	yes (N/A/V)		root-level affixation
(ii)	Imposes input morpho-syntactic features	sometimes, not always		categorization
(iii)	Can appear above inflection	no		word-level modification
(iv)	Can appear above derivation	hardly	sometimes	word-level affixation
(v)	Productive	not so much	fairly	anything but
(vi)	Semantically compositional	mostly no if features not preserved mostly yes if features preserved		root-level categorization

TABLE 3: SUMMARY OF THE BEHAVIOR OF *-et* AND *-ette* WITH RESPECT TO THE DIAGNOSTICS DEFINED IN SECTION 1.2.

Let us call DIM the abstract representation of the diminutive feature. Row (ii) of Table 3 suggests DIM can be a modifier,[18] but also sometimes a head, when features such as gender are not preserved. Row (i) moreover suggests that some derivations should be root-level, and row (vi) adds that such derivations are at least sometimes of the head-categorization kind, due to the correlation between featural and semantic opacity. We take the differential behavior of *-et* vs. *-ette* in rows (iv) and (v) to be evidence that *-et* is only compatible

---

[18] We can say that because as we will see in the next section, morpho-syntactic features such as gender are *most of the time* preserved, therefore it is unlikely that feature preservation results from repeated, accidental, gender-matching *head* categorization.

with root-level derivations while *-ette* is compatible with at least some word-level derivations. In brief, we think these data suggest that DIM is ambiguous between a root-level modifier following the schema in (14a) and two heads, one with masculine features merged at the root-level, and one with feminine features merged at the word-level, as schematized in (14b).

(14)   Insertion Rules for *-et* and *-ette* (first attempt)

a. $[\text{DIM}^{rl}_{\text{mod}}]$ $\leftrightarrow$ *-et*   / [M] __
         -ette   elsewhere

b. $[\text{DIM}^{rl}_{\text{M}}]$ $\leftrightarrow$ *-et*
      $[\text{DIM}^{wl}_{\text{F}}]$ $\leftrightarrow$ *-ette*

We assumed here that the modifiers and heads compatible with *-et* and *-ette* were all of the DIM kind, but nothing in principle prevents those elements to actually have different formal features and a different semantics (we will see in the next Section that this might actually be the case). Before going further into the formal analysis, let us try to verify the following coarse predictions. If *-et* and *-ette* are indeed disentangled *via* two distinct heads with different properties, then whatever those heads mean, their differences in attachment height make us expect frequency and meaning asymmetries to arise between *-et* and *-ette* in the French lexicon. The next section thus presents a quantitative analysis of the distribution and semantics of French diminutive nominals in *-et/-ette*.

3. QUANTITATIVE ANALYSIS OF NOMINALS IN *-ET(TE)*

3.1 Methodology

From a list of French words provided by the Ubuntu Linux Operating System (346,200 entries), we collected all nouns ending in *-et* or *-ette*. We checked that the words we retained were nouns using the spaCy part-of-speech tagger for French based on the `fr_core_news_md` pipeline (Honnibal & Montani 2017). We then generated, for each noun in *-et/-ette*, an approximation of its base by truncating the suffix *-et/-ette*. For each approximate base, we checked if it ended up being "close enough" to an existing nominal of the lexicon – if yes, the resulting base-diminutive pair was added to our tentative dataset. We took a nominal from the lexicon to be "close enough" to an approximate base of length ℓ if (i) the length-$(\ell - 1)$ prefix[20] of the nominal and that of the approximate

---

[20] The choice of a length-$(\ell - 1)$ prefix (as opposed to a length-ℓ prefix which would have covered the entire approximate base) was intended to retain base-diminutive pairs whereby the

base were the same; and (ii) the total length of the nominal did not exceed that of the diminutive (i.e. $\ell + 4$ is the diminutive ended in *-ette*, and $\ell + 2$ if it ended in *-et*).

Further manual filtering of this dataset involved checking if each pair made sense[21] using online resources (Larousse online dictionary, Wiktionary) and introspection, and checking once again if the forms were purely nominal.[22] Lastly, the dataset was supplemented by pairs generated *via* introspection (not all of them being documented in dictionaries) – for a total of 262 nouns in *-ette* and 146 nouns in *-et*. Further statistics are compiled in Table 4.

Derivative → Base ↓	Feminine in *-ette*	Masculine in *-et*	Total # of derivatives	% of bases with both suffixes
Feminine (233 bases)	218	47	265	$^{32}/_{233} = 13.7\%$
Masculine (133 bases)	44	99	143	$^{10}/_{133} = 7.5\%$
Total	262	146	408	$^{42}/_{366} = 11.5\%$

TABLE 4: DATASET STATISTICS; ▩=GENDER-PRESERVING SUFFIXATION; ▩=F-TO-M MISMATCHES; ▩=M-TO-F MISMATCHES. NOTE THAT CERTAIN BASES WERE COMPATIBLE WITH BOTH *-et* AND *-ette*; THE LAST COLUMNS SUMMARIZES HOW FREQUENTLY THIS WAS.

3.2 Distribution of *-ette* vs. *-et* suffixation and gender mismatches

Regardless of the gender of the base, *-ette* suffixation was around 2.2 times more frequent than *-et* suffixation ($^{218}/_{99} \sim 2.2$), according to Table 4 (Observation I). This frequency asymmetry is probably not due to the baserates of feminine vs. masculine nominals in the initial lexicon we used: on a sample of 3,782 nouns (identified as such by spaCy on a random sample of 10,000 words from the lexicon), 412 were identified as feminine and 454 as masculine, using grammatical gender data scraped from the French Wiktionary *via* the Python WiktionaryParser module.[23] The distribution of feminine vs. masculine

---

diminutized base underwent some orthographical/phonological changes triggered by suffixation. For instance, in *bateau/batelet*, which is a well-formed base-derivative pair, the approximate base obtained by truncating *-et* is *batel*, which is of length 5. And the length 4 prefix of *bateau* is *bate* which turns out to be identical to the length 4 prefix of *batel*, so *bateau* and *batelet* are correctly predicted to be "close enough". This would not have been the case if, instead of a prefix of length $\ell - 1$, we had chosen to check a prefix of length ℓ: *batel* differs from *batea*.

[21] In particular, the pair had to feature words morphologically related *in French*, and not, e.g. an English base already suffixed with *-ette* in English and adopted as is by French (e.g. *majorette*).

[22] Cases in which it was unclear whether the base was nominal, which happens, for instance, when the nominal is ambiguous with a deverbal, as in *déjeuner* ('(have) lunch'), were excluded.

[23] This module allowed us to access the content of the Wiktionary page of words from our lexicon, formatted as a JSON object. To access the grammatical gender of nominals, we focused on the section of the JSON associated with the part-of-speech "noun" (if present), and retrieved the

nominals in our lexicon was thus approximately even (47.6% feminine, 52.4% masculine), and consistent with previous lexicographic analyses (Roché 1992). The observed asymmetry therefore implies that *-ette* is overall more productive than *-et*, which would be surprising if both were realizations of the same underlying morpheme.

Additionally, the proportion of gender-mismatches appears higher for M-bases (M-to-F mismatch) than F-bases (F-to-M mismatch), as per a two-tailed Z-score for 2 population proportions (Observation II):

$$\hat{\mathbb{P}}[\text{-}et|\text{F-base}] = {}^{47}/_{233} \sim 20\% <_{p=.006} \hat{\mathbb{P}}[\text{-}ette|\text{M-base}] = {}^{44}/_{133} \sim 33\%$$

The amplitude of this discrepancy is approximately the same as the one recorded for *-et*/*-ette* forms in general (${}^{33}/_{20}\sim 1.8$, Observation III). This correlation would be unexpected if gender-mismatching forms were morphologically unrelated to gender-matching forms, e.g. if they were pure lexicalizations as the previous literature assumed (Dauzat 1937; Milner 1989 i.a.).

3.3 Diminutive semantics

Table 5 summarizes the proportion of "true" diminutives per condition. The diminutive character of each derivative was assessed by introspection using the "redundancy test" introduced below, and by consulting additional online resources (Larousse dictionary, Wiktionary) on a case-by-case basis.

(15) *Redundancy test (Milner 1989):* for any given base *B*, if the phrase "small *B-et(te)*" does *not* sound redundant, then *B-et(te)* is *not* purely diminutive.

The rationale behind the above test is that it appears redundant to describe something as small using an adjective, when this information is already conveyed by a morpheme attached to the noun, as it is the case in (16). However, when the diminutive used does not convey a difference in size but instead creates a new word, the feeling of redundancy goes away, as in (17).

---

gender encoded in this entry as a letter ("m", "f", or very rarely "?"). If no JSON was produced for the target word (meaning: the word was not listed in the Wiktionary), or if no or more than one gender label was associated with it, the word was ignored. Out of the 3,782 putative nominals tagged with spaCy, 2,916 were ignored, 2,874 because no gender label was found, most if not all of the time due to the absence of a "noun" section in the Wiktionary definition (meaning: the word was probably misclassified as a noun by spaCy); 41 because the word was not listed in the Wiktionary, and 1 because multiple genders were recorded.

(16) a. *??une petite maisonette*
      ??a   small small-house

    b. *??un petit balconnet*
        ??a small small-balcony

    c. *??un petit jardinet.*
        ??a small small-garden

(17) a. *une petite charette*
      a   small cart

    b. *un petit boulet*
        a small (cannon)ball

    c. *une petite cigarette.*
        a small cigarette

Derivative → Base ↓	Feminine in *-ette*	Masculine in *-et*	Total
Feminine	$161/218 = 73.9\%$	$12/47 = 25.5\%$	$173/265 = 65.3\%$
Masculine	$15/44 = 34.1\%$	$60/99 = 60.6\%$	$75/143 = 52.4\%$
Total	$176/262 = 67.2\%$	$72/146 = 49.3\%$	$248/408 = 60.8\%$

TABLE 5: PROPORTIONS OF "TRUE" DIMINUTIVES PER GROUP. SAME COLOR-CODING AS BEFORE.

From Table 5 it can be seen that more gender-matching forms than gender-mismatching ones appear to have a true diminutive semantics, as per a two-tailed Z-score for 2 population proportions (Observation IV).

$$\hat{\mathbb{P}}[\text{DIM}|\text{F-base-}ette \text{ or M-base-}et] = 161+60/218+99 \sim 70\%$$

$$>_{p<.00001}$$

$$\hat{\mathbb{P}}[\text{DIM}|\text{F-base-}et \text{ or M-base-}ette] = 12+15/47+44 \sim 30\%$$

This is in line with Milner's observation about the semantic effects of gender-mismatch. However, an asymmetry driven by the gender of the suffix arises in both the "match" and "mismatch" conditions: non-mismatching F-forms in *-ette* are more likely to be diminutive than non-mismatching forms in *-et*, as per a two-tailed Z-score for 2 population proportions (Observation V).

$$\hat{\mathbb{P}}[\text{DIM}|\text{F-base-}ette] = 161/218 \sim 74\% >_{p=.02} \hat{\mathbb{P}}[\text{DIM}|\text{M-base-}et] = 60/99 \sim 60\%$$

The same patterns holds for mismatching forms (although non-significant, potentially due to small sample sizes; Observation VI).[25] It is worth noting that even if the gender-matching forms are more likely than the mismatching ones to exhibit a true diminutive semantics, the proportions of "true" diminutives in gender-matching conditions are not at ceiling. Gender-matching pairs like those in (18) and (19) exemplify, *for both genders*, some kind of loose, shape-related semantic relationship between the base and the derivative.

---

[25] $\hat{\mathbb{P}}[\text{DIM}|\text{M-base-}ette] = 15/44 \sim 34\% > \hat{\mathbb{P}}[\text{DIM}|\text{F-base-}et] = 12/47 \sim 26\%$

(18) a. $oeil_M \rightarrow (oeill\text{-}et)_M$
          'eye' $\rightarrow$ 'eyelet'

      b. $arc_M \quad\quad\quad \rightarrow (arch\text{-}et)_M$
          'bow (archery)' $\rightarrow$ 'bow (music)'

(19) a. $barre_F \quad\quad\quad \rightarrow (barr\text{-}ette)_F$
          'bar (construction)' $\rightarrow$ 'hair-clip'

      b. $coquille_F \rightarrow (coquill\text{-}ette)_F$
          'shell' $\quad \rightarrow$ 'elbow pasta'

These data seem to mean that mismatches *per se* cannot explain all the semantic variation there is in seemingly diminutive forms: something about the suffixes themselves, and how they combine with their base, must be at play. In the next Section, we revise our preliminary account of *-et* vs. *-ette*, and show how it captures these data as well as most of the patterns from Section 2.

4. FORMAL ANALYSIS

4.1 Novel claims

As we hinted at the end of Section 2, we want to propose that *-ette* is ambiguous between a transparent root-level modifier $\text{DIM}^{rl}_{\text{mod}}$ and a feminine-marked, word-level head. We now want to claim that the semantics of this head, that we call $\text{SIM}^{wl}_{\text{F}}$, is not purely diminutive, but instead indicates similarity with respect to a salient feature (shape typically).[27] Likewise, we propose that *-et* is ambiguous between the transparent root-level modifier $\text{DIM}^{rl}_{\text{mod}}$ and a masculine-marked, root-level version of the SIM-head, $\text{SIM}^{rl}_{\text{M}}$. This is schematized in (20). Figure 4 illustrates the three possible structures compatible with *-et/-ette* suffixation.

(20) Insertion Rules for *-et* and *-ette* (final attempt)

   a. $\begin{aligned}[\text{DIM}^{rl}_{\text{mod}}] \quad &\leftrightarrow \quad \textit{-et} \quad\quad / [\text{M}] \, \_\_ \\ & \textit{-ette} \quad \text{elsewhere}\end{aligned}$

   b. $\begin{aligned}[\text{SIM}^{rl}_{\text{M}}] \quad &\leftrightarrow \quad \textit{-et} \\ [\text{SIM}^{wl}_{\text{F}}] \quad &\leftrightarrow \quad \textit{-ette}\end{aligned}$

---

[27] Milner (1989) and Delhay (1999) had made a similar observation to describe cases of gender-mismatch. We will see that our analysis allows to explain why empirically, this observation partly extends to cases where gender is preserved.
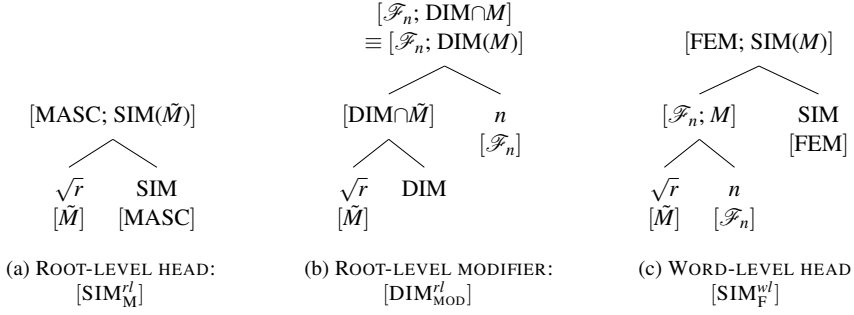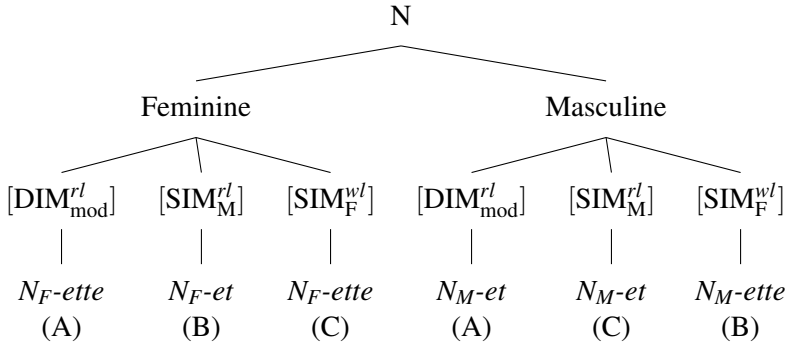
$$[\mathscr{F}_n; \mathrm{DIM} \cap M]$$
$$\equiv [\mathscr{F}_n; \mathrm{DIM}(M)]$$

$$[\mathrm{MASC}; \mathrm{SIM}(\tilde{M})]$$
$$[\mathrm{DIM} \cap \tilde{M}] \quad n$$
$$[\mathscr{F}_n]$$

$$[\mathrm{FEM}; \mathrm{SIM}(M)]$$

$$[\mathscr{F}_n; M] \quad \mathrm{SIM}$$
$$[\mathrm{FEM}]$$

$$\sqrt{r} \quad \mathrm{SIM}$$
$$[\tilde{M}] \quad [\mathrm{MASC}]$$

$$\sqrt{r} \quad \mathrm{DIM}$$
$$[\tilde{M}]$$

$$\sqrt{r} \quad n$$
$$[\tilde{M}] \quad [\mathscr{F}_n]$$

(a) ROOT-LEVEL HEAD:
$$[\mathrm{SIM}_{\mathrm{M}}^{rl}]$$

(b) ROOT-LEVEL MODIFIER:
$$[\mathrm{DIM}_{\mathrm{MOD}}^{rl}]$$

(c) WORD-LEVEL HEAD
$$[\mathrm{SIM}_{\mathrm{F}}^{wl}]$$

FIGURE 4: SUMMARY OF THE 3 POSSIBLE STRUCTURES ASSOCIATED WITH -*et*/-*ette*, AND OF THEIR CONSEQUENCES ON THE FEATURES AND MEANING OF THE OUTPUT WORD.

Lastly, the decision tree in (21) describes the possible derivations of the various surface forms of a nominal N depending on its gender, given (20). Within this decision tree, three patterns emerge. In terms of form, Pattern (A) corresponds to gender-matching, purely DIM suffixation. Pattern (B) corresponds to gender-mismatching SIM suffixation. Finally, pattern (C) corresponds to *accidentally* gender-matching SIM suffixation ("accidental" because SIM still imposes its own gender which just happens to match that of the base). In terms of semantics, Pattern (A) is supposed to yield the most productive and transparent diminutive semantics. Pattern (B) is expected to yield a non-productive, doubly non-diminutive semantics (due to the posited semantics for SIM, and to the non-compositional effect of root-level categorization). Pattern (C) finally, is expected to yield a productive, compositional, and similarity-based semantics (that in some cases might be accidentally diminutive).

(21)   *Possible surface suffixations of noun N depending on its gender and the abstract suffix used.*

N

Feminine

Masculine

$$[\mathrm{DIM}_{\mathrm{mod}}^{rl}] \quad [\mathrm{SIM}_{\mathrm{M}}^{rl}] \quad [\mathrm{SIM}_{\mathrm{F}}^{wl}] \quad [\mathrm{DIM}_{\mathrm{mod}}^{rl}] \quad [\mathrm{SIM}_{\mathrm{M}}^{rl}] \quad [\mathrm{SIM}_{\mathrm{F}}^{wl}]$$

$$N_F\text{-}ette \quad N_F\text{-}et \quad N_F\text{-}ette \quad N_M\text{-}et \quad N_M\text{-}et \quad N_M\text{-}ette$$
(A)   (B)   (C)   (A)   (C)   (B)

4.2 Verifying the quantitative observations of Section 3

Let us start with the observations pertaining to the relative frequencies of *-et* vs. *-ette* forms. We make the hypothesis that the probability for $[\text{DIM}_{\text{mod}}^{rl}]$, $[\text{SIM}_{\text{M}}^{rl}]$, or $[\text{SIM}_{\text{F}}^{wl}]$ to surface is independent of the gender of the base they apply to, and that, moreover, $[\text{SIM}_{\text{M}}^{rl}]$ being a root-level head, it is less productive and thus less likely to be inserted than $[\text{DIM}_{\text{mod}}^{rl}]$ and $[\text{SIM}_{\text{F}}^{wl}]$. We first derive the fact that *-ette* forms are overall more frequent than *-et* forms (Observation I): *-ette* forms result from $[\text{DIM}_{\text{mod}}^{rl}]$ or $[\text{SIM}_{\text{F}}^{wl}]$ affixation if the base is feminine, and $[\text{SIM}_{\text{F}}^{wl}]$ affixation if it is masculine, while *-et* forms result from $[\text{DIM}_{\text{mod}}^{rl}]$ or $[\text{SIM}_{\text{M}}^{rl}]$ affixation if the base is masculine, and $[\text{SIM}_{\text{M}}^{rl}]$ affixation if it is feminine. Canceling out $[\text{DIM}_{\text{mod}}^{rl}]$ from the equation, the target inequality then follows from the fact $[\text{SIM}_{\text{F}}^{wl}]$ is more frequent than $[\text{SIM}_{\text{M}}^{rl}]$. We can also show that M-to-F mismatches are more frequent than F-to-M mismatches (Observation II): M-to-F mismatches are the result of $[\text{SIM}_{\text{F}}^{wl}]$ categorization while F-to-M mismatches are the result of $[\text{SIM}_{\text{M}}^{rl}]$ categorization, so the prediction immediately follows. Moreover, the amplitude of this frequency difference is accurately predicted to match that of the frequency difference between *-et* and *-ette* in the general case (Observation III).

The account also covers our quantitative observations pertaining to the diminutive semantics of the gender-matching vs. mismatching forms. First, the observation that more gender-matching forms than gender-mismatching ones appear to have a true diminutive semantics (Observation IV) is captured. To see that, notice that gender-matching forms correspond to Patterns (A) and (C) (for both genders) which themselves correspond to the application of $[\text{DIM}_{\text{mod}}^{rl}]$ (two occurrences in Tree (21)), $[\text{SIM}_{\text{F}}^{wl}]$ and $[\text{SIM}_{\text{M}}^{rl}]$ (one occurrence each, depending on gender), while gender-mismatching forms correspond to Patterns (B) which corresponds to $[\text{SIM}_{\text{M}}^{rl}]$ or $[\text{SIM}_{\text{F}}^{wl}]$ (depending on gender). So, canceling out the cases of head categorization ($[\text{SIM}_{\text{M}}^{rl}]$/$[\text{SIM}_{\text{F}}^{wl}]$) from the equation, we see that gender-matching forms, unlike gender-mismatching ones, comprise forms resulting form the the application of $[\text{DIM}_{\text{mod}}^{rl}]$ which is purely diminutive.

Second, we capture the observation that non-mismatching F-forms in *-ette* are more diminutive than non-mismatching M-forms in *-et* – and that both are still fairly diminutive overall (Observation V). Non-mismatching F-forms in *-ette* come from $[\text{DIM}_{\text{mod}}^{rl}]$ or $[\text{SIM}_{\text{F}}^{wl}]$, while non-mismatching M-forms in *-et* come from $[\text{DIM}_{\text{mod}}^{rl}]$ or $[\text{SIM}_{\text{M}}^{rl}]$. Canceling out $[\text{DIM}_{\text{mod}}^{rl}]$ from the equation, the asymmetry can be explained by the difference in semantic compositionality between $[\text{SIM}_{\text{F}}^{wl}]$ and $[\text{SIM}_{\text{M}}^{rl}]$. Indeed, even if both suffixes express base-derivative *similarity*, they might accidentally express diminutization (after all, a

kind of similarity relation); those instances of accidental diminutization will be fully semantically compositional if $[\text{SIM}_\text{F}^{wl}]$ is involved (word-level derivation), but not so if $[\text{SIM}_\text{M}^{rl}]$ is involved (root-level derivation). In other words, $[\text{SIM}_\text{F}^{rl}]$ is more likely to lead to a diminutive semantics (even if it is rare) than $[\text{SIM}_\text{F}^{wl}]$. Beyond the asymmetry, the fact *both* forms (F-forms in *-ette* and M-forms in *-et*) are compatible with $[\text{DIM}_\text{mod}^{rl}]$ explains why both are still fairly diminutive overall.

Third, this reasoning directly extends to the contrast between mismatching M-forms in *-ette* (more diminutive) vs. mismatching F-forms in *-et* (less diminutive), that however did not reach significance in our quantitative analysis (Observation VI). M-forms in *-ette* come from $[\text{SIM}_\text{F}^{wl}]$, while F-forms in *-et* come from $[\text{SIM}_\text{M}^{rl}]$, directly giving the expected asymmetry and also at the same time, predicting the overall low counts of pure diminutives for those forms. As a final note for this Section, the qualitative observations in (18) and (19) are captured by the fact that such forms might be the result of Pattern (C) that we dubbed "accidental" gender-matching: within that pattern, the gender of the suffix matches that of the base, yet the underlying morpheme is not DIM, but rather, a non-purely diminutive SIM-head whose gender accidentally matches the gender of the base.

4.3 Verifying the qualitative observations of Section 2

To close this paper let us go back to the broader observations from Section 2 and see if our analysis of *-et/-ette* extends to them. In Section 2.2 we observed that *-ette* is more productive than *-et* when applied to recent loanwords from English. The observed *contrast* is predicted by our account given that gender-matching diminutized loanwords are compatible with both Pattern (A) (transparent modification) and Pattern (C) (accidentally gender-matching categorization), the latter being the source of the productivity asymmetry between *-et* and *-ette*, given that $[\text{SIM}_\text{M}^{rl}]$, which is root-level, is expected to be more productive than $[\text{SIM}_\text{F}^{wl}]$, which is word-level. However, the fact that productivity is overall quite low on loanwords (especially in the case of *-et*) remains mysterious given the availability of Pattern (A) (productive) for both genders.

In Section 2.3, we made the observation that *-et/-ette* surfaces on adjectives (in a featurally transparent way) and verbs (with mostly compositional *-ette* markings), in addition to nouns. We think that both cases can be captured by $[\text{DIM}_\text{mod}^{rl}]$, assuming that *-ette* on verbs is the elsewhere realization of this suffix in the absence of any gender marking (cf. the Insertion Rule in (20a)). The only remaining issue lies in the low degree of productivity of adjective-to-adjective *-et(te)* suffixation, that is unexpected if $[\text{DIM}_\text{mod}^{rl}]$ is available.

In Section 2.4 we observed that *-ette* could appear above derivational morphology, while it was hard to find similar examples involving *-et*. First, it is difficult to guarantee that this fact is solely due to issues with morphological derivation, and not maybe just a consequence of the fact that more derived words that are easily diminutizable (e.g., nouns in *-tion*) happen to be consistently feminine. But if the latter explanation could be ruled out, our account would explain the observed asymmetry by the fact that derivations where *-et* surfaces are necessarily root-level, while derivations where *-ette can* be word-level (in cases of categorization). As already discussed, the impossibility to stack diminutive suffixes in French is perhaps the most problematic, especially given our account that allows high and low suffixes. We sketched a potential explanation in terms of blocking back in Section 2.4, that can be maintained under our current analysis.

5. CONCLUSION AND OUTLOOK

Building on (Wiltschko & Steriopolo 2007; Steriopolo 2017), we argued that the quantitatively observed differences in productivity and semantic transparency between *-ette* and *-et* could be explained by a two-way ambiguity. *-ette* was stipulated to be ambiguous between a root-level DIM modifier and a word-level, feminine-marked SIM head, expressing a (typically shape-based) form of similarity between the base and its derivative. *-et* on the other hand, was stipulated to be ambiguous between the root-level DIM modifier (same as for *-ette*) and a root-level, masculine-marked SIM head. This makes French similar to languages like Russian, which was argued to possess both root-level head diminutives and word-level modifier diminutives (although the two kinds were realized by totally different morphemes).

Crucially, our account provided a morphosyntactic explanation as to why gender-mismatches correlate with some form of *semantic* mismatch, following insights from previous work on diminutives (De Belder *et al.* 2014) and beyond (Arad 2003). It would be interesting to extend that kind of semantics-oriented approach to morphology to other languages featuring several diminutive or augmentative suffixes with differing morphological properties in terms of attachment height and attachment type – typically, Russian. Other perspectives for future work may include extending this kind of analysis of diminutive suffixes to other pairs of French suffixes with gender marking (e.g. the suffix *-ot(te)* in (22)).

(22) a.   *cage* → *(cage-ot)$_M$*   c.   *cul*   → *(cull-otte)$_F$*
               'cage' → 'crate'   'buttocks' → 'panties'

       b.   *char*   → *(charri-ot)$_M$*   d.   *bouger* → *(bouge-otte)$_F$*
               'chariot' → 'cart'   'move' → 'restlessness'

REFERENCES

Acquaviva, P. (2009). Roots and lexicality in Distributed Morphology. In *York Essex Morphology Meeting 2*. 1–21.

Arad, M. (2003). Locality constraints on the interpretations of roots. *Natural Language and Linguistic Theory*, 21(4). 737–778. `https://doi.org/10.1023/a:1025533719905`.

Arad, M. (2005). *Roots and Patterns: Hebrew Morpho-syntax*. Dordrecht: Springer.

Aronoff, M. (1976). *Word Formation in Generative Grammar*.

Aronoff, M. & M. Lindsay (2014). Productivity, Blocking, and Lexicalization. In *The Oxford Handbook of Derivational Morphology*. Oxford University Press. `https://doi.org/10.1093/oxfordhb/9780199641642.013.0005`.

Bally, C. (1932). *Linguistique générale et linguistique française*.

Borer, H. (2005). *Structuring sense. Volume 1: In name only*. Oxford: OUP.

Borer, H. (2013). *Structuring sense. Volume 1: Taking form*. Oxford: OUP.

Dal, G. (1997). *Grammaire du suffixe -et(te)*. Ph.D. thesis, Paris.

Dauzat, A. (1937). L'appauvrissement de la dérivation française. *Le Français moderne*, 4. 289–299.

De Belder, M., N. Faust & N. Lampitelli (2014). *On an inflectional and a derivational diminutive*. Oxford University Press, 149–163.

Deal, A.R. (2016). Plural exponence in the Nez Perce DP: a DM analysis. *Morphology*, 26(3-4). 313–339. `https://doi.org/10.1007/s11525-015-9277-9`.

Delhay, C. (1995). Le diminutif: la diminution sans comparaison? *Faits de langues*, 3(5). 63–72. `http://dx.doi.org/10.3406/flang.1995.976`.

Delhay, C. (1999). "Diminutifs" et niveaux de catégorisation. *Faits de langues*, 7(14). 79–87. `https://doi.org/10.3406/flang.1999.1268`.

Fathi, R. & J. Lowenstamm (2016). The gender assignment pattern of French nouns: Selection and allomorphy. *Morphology*, 26(3–4). 477–509. `http://dx.doi.org/10.1007/s11525-016-9287-2`.

Ferrari, F. (2005). *A syntactic analysis of the nominal systems of Italian and Luganda: how nouns can be formed in the syntax*. Ph.D. thesis, New York University.

Fradin, B. (2003). Le traitement de la suffixation en -et. *Langages*, 37(152). 51–77. `http://dx.doi.org/10.3406/lgge.2003.2439`.

Halle, M. & A. Marantz (1993). Distributed Morphology and the Pieces of Inflection. In K. Hale & S.J. Keyser (eds.) *The View from Building 20*. Cambridge, MA: MIT Press, 111–176.

Halle, M. & A. Marantz (1994). Some Key Features of Distributed Morphology. In *Papers on Phonology and Morphology*, 275–288. MITWPL 21, Cambridge, MA: MIT Working Papers in Linguistics.

Harley, H. & R. Noyer (1998). Licensing in the non-lexicalist lexicon: Nominalizations, vocabulary items and the encyclopedia. In H. Harley (ed.) *MIT Working Papers in Linguistics 32: Papers from the UPenn/MIT roundtable on argument structure and aspect*, 119–137. Cambridge, MA: MITWPL.

Hasselrot, B. (1957). Etudes sur la formation diminutive dans les langues romanes. In *Acta Universitatis Upsaliensis 1957 : 11*. Uppsala: Lundequistska Bokhandeln.

Hasselrot, B. (1972). Etude sur la vitalité de la formation diminutive française au xxe siècle. In *Acta Universitatis Upsaliensis, Studia Romanica Upsaliensia 8*. Uppsala: Almqvist & Wiksells.

Honnibal, M. & I. Montani (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear.

Jurafsky, D. (1996). Universal Tendencies in the Semantics of the Diminutive. *Language*, 72(3). 533. https://doi.org/10.2307/416278.

Kihm, A. (2005). *Noun class, gender and the lexicon-syntax-morphology interfaces: A comparative study of Niger-Congo and Romance languages*. Oxford: OUP, 459–512.

Kramer, R. (2015). *The Morphosyntax of Gender*. Oxford Studies in Theoretical Linguistics 58, Oxford: Oxford University Press.

Kramer, R. (2016). The location of gender features in the syntax. *Language and Linguistics Compass*, 10(11). 661–677. https://doi.org/10.1111/lnc3.12226.

Lecarme, J. (2002). *Gender "polarity": Theoretical aspects of Somali nominal morphology*. Somerville, Massachusetts: Cascadilla Press, 109–141.

Marantz, A. (1997). No escape from syntax. *University of Pennsylvania Working Papers in Linguistics*, 4. 201–225.

Marantz, A. (2001). *Words*.

Milner, J.C. (1989). Genre et dimension dans les diminutifs français. *Linx*, 21(1). 191–201. https://doi.org/10.3406/linx.1989.1141.

Roché, M. (1992). Le masculin est-il plus productif que le féminin ? *Langue française*, 96. 113–124.

Scalise, S. (1986). *Generative morphology, 2nd edn*. Dordrecht: Foris.

Steriopolo, O. (2017). Syntactic variation in diminutive suffixes: Russian, kolyma yukaghir, and itelmen. *Languages*, 2(4). https://www.mdpi.com/2226-471X/2/4/23.

Steriopolo, O., G. Markopoulos & V. Spyropoulos (2023). *Parameters of variation in the syntax of expressive suffixes: Case studies of Russian, German, Spanish and Greek*. Cambridge University Press, 361–392.

Stump, G.T. (1993). How peculiar is evaluative morphology? *Journal of Linguistics*, 29(1). 1–36. http://dx.doi.org/10.1017/S0022226700000037.

Weber, M. (1963). *Contributions à l'étude du diminutif en français moderne : essai de systématisation*. Ph.D. thesis.

Wiltschko, M. & O. Steriopolo (2007). Parameters of variation in the syntax of diminutives. In M. Radisic (ed.) *Proceedings of the 2007 Canadian Linguistics Association Annual Conference*.

Adèle Hénot-Mortier
Massachusetts Institute of Technology
32 Vassar Street - Cambridge, MA
United States
e-mail: adele.mortier@gmail.com
https://orcid.org/0000-0003-3173-9580