Morpho-semantics of the French diminutive suffix -et(te)[1]

Adèle Hénot-Mortier (MIT)

September 15, 2023

International Symposium of Morphology 2023

Background

- French assigns grammatical gender (**M**asculine or **F**eminine) to nominals and is endowed with a quite productive "diminutive" suffix **-et**/**-ette**.

(1)   a.   maison$_{\textbf{F}}$ → (maisonn-***ette***)$_{\textbf{F}}$
             'house' → 'small (cute) house'

    b.   balcon$_{\textbf{M}}$ → (balconn-***et***)$_{\textbf{M}}$
             'balcony' → 'small (cute) balcony'

- Because **M**-bases are often affixed with the **M**-variant of the diminutive (**-et**) and **F**-bases with the **F**-variant (**-ette**), traditional grammars implicitly assumed that **-et** and **-ette** were allomorphs dependent on the gender features of the base.
- A correlate of this assumption is that **-et** and **-ette** should be linked to the exact same diminutive semantics.

Background

- French assigns grammatical gender (**M**asculine or **F**eminine) to nominals and is endowed with a quite productive "diminutive" suffix **-et**/**-ette**.

(1) a. $\text{maison}_{\mathbf{F}} \rightarrow (\text{maisonn-}\mathbf{\textit{ette}})_{\mathbf{F}}$
         'house' $\rightarrow$ 'small (cute) house'

     b. $\text{balcon}_{\mathbf{M}} \rightarrow (\text{balconn-}\mathbf{\textit{et}})_{\mathbf{M}}$
         'balcony' $\rightarrow$ 'small (cute) balcony'

- Because **M**-bases are often affixed with the **M**-variant of the diminutive (**-et**) and **F**-bases with the **F**-variant (**-ette**), traditional grammars implicitly assumed that **-et** and **-ette** were allomorphs dependent on the gender features of the base.

- A correlate of this assumption is that **-et** and **-ette** should be linked to the exact same diminutive semantics.

2

Background

- French assigns grammatical gender (**M**asculine or **F**eminine) to nominals and is endowed with a quite productive "diminutive" suffix **-et**/**-ette**.

(1) a.   maison$_\mathbf{F}$ → (maisonn-***ette***)$_\mathbf{F}$
         'house' → 'small (cute) house'

     b.   balcon$_\mathbf{M}$ → (balconn-***et***)$_\mathbf{M}$
         'balcony' → 'small (cute) balcony'

- Because **M**-bases are often affixed with the **M**-variant of the diminutive (**-et**) and **F**-bases with the **F**-variant (**-ette**), traditional grammars implicitly assumed that **-et** and **-ette** were allomorphs dependent on the gender features of the base.

- A correlate of this assumption is that **-et** and **-ette** should be linked to the exact same diminutive semantics.

Background

- French assigns grammatical gender (**M**asculine or **F**eminine) to nominals and is endowed with a quite productive "diminutive" suffix **-et**/**-ette**.

(1)   a.   maison$_\mathbf{F}$ $\rightarrow$ (maisonn-**ette**)$_\mathbf{F}$
           'house' $\rightarrow$ 'small (cute) house'

    b.   balcon$_\mathbf{M}$ $\rightarrow$ (balconn-**et**)$_\mathbf{M}$
           'balcony' $\rightarrow$ 'small (cute) balcony'

- Because **M**-bases are often affixed with the **M**-variant of the diminutive (**-et**) and **F**-bases with the **F**-variant (**-ette**), traditional grammars implicitly assumed that **-et** and **-ette** were allomorphs dependent on the gender features of the base.

- A correlate of this assumption is that -et and -ette should be linked to the exact same diminutive semantics.

Background

- French assigns grammatical gender (**M**asculine or **F**eminine) to nominals and is endowed with a quite productive "diminutive" suffix **-et**/**-ette**.

(1)   a.   maison$_\mathbf{F}$ $\rightarrow$ (maisonn-***ette***)$_\mathbf{F}$
             'house' $\rightarrow$ 'small (cute) house'

       b.   balcon$_\mathbf{M}$ $\rightarrow$ (balconn-***et***)$_\mathbf{M}$
             'balcony' $\rightarrow$ 'small (cute) balcony'

- Because **M**-bases are often affixed with the **M**-variant of the diminutive (**-et**) and **F**-bases with the **F**-variant (**-ette**), traditional grammars implicitly assumed that **-et** and **-ette** were allomorphs dependent on the gender features of the base.
- A correlate of this assumption is that **-et** and **-ette** should be linked to the exact same diminutive semantics.

- Milner (1989) however observed that **-ette** may attach to **M**-bases and **-et** to **F**-bases – a phenomenon we dub **gender-mismatch** – leading to a looser semantic relationship between the base and the derived form.

(2)   char$_M$ $\longrightarrow$ (char-*ette*)$_F$ (3)   a.   boule$_F$ $\longrightarrow$ (boul-*et*)$_M$
   char$_M$ *$\longrightarrow$ (char-*et*)$_M$ 'ball' $\longrightarrow$ 'cannonball'
   'chariot' $\longrightarrow$ 'cart'

             b.   boule$_F$ $\longrightarrow$ (boul-*ette*)$_F$
              'ball' $\longrightarrow$ 'small ball'

- The pairs in (2) and (3a) would be unexpected if the suffix simply agreed in gender with the base: rather, it seems that in at least certain cases, the suffix introduces its own gender (a phenomenon documented in other languages, cf. Kramer, 2015).

3

- Milner (1989) however observed that **-ette** may attach to **M**-bases and **-et** to **F**-bases – a phenomenon we dub **gender-mismatch** – leading to a looser semantic relationship between the base and the derived form.

(2)   char$_\mathbf{M}$ $\longrightarrow$ (char-***ette***)$_\mathbf{F}$ (3)   a.   boule$_\mathbf{F}$ $\longrightarrow$ (boul-***et***)$_\mathbf{M}$

         char$_\mathbf{M}$ *$\longrightarrow$ (char-***et***)$_\mathbf{M}$ 'ball' $\longrightarrow$ 'cannonball'

         'chariot' $\longrightarrow$ 'cart'

                                             b.   boule$_\mathbf{F}$ $\longrightarrow$ (boul-***ette***)$_\mathbf{F}$

                                                   'ball' $\longrightarrow$ 'small ball'

- The pairs in (2) and (3a) would be unexpected if the suffix simply agreed in gender with the base: rather, it seems that in at least certain cases, the suffix introduces its own gender (a phenomenon documented in other languages, cf. Kramer, 2015).

3

A more intricate picture (Milner, 1989)

- Milner (1989) however observed that **-ette** may attach to **M**-bases and **-et** to **F**-bases – a phenomenon we dub **gender-mismatch** – leading to a looser semantic relationship between the base and the derived form.

(2) char$_\mathbf{M}$ $\longrightarrow$ (char-**ette**)$_\mathbf{F}$
     char$_\mathbf{M}$ $\overset{*}{\longrightarrow}$ (char-**et**)$_\mathbf{M}$
     'chariot' $\longrightarrow$ 'cart'

(3) a. boule$_\mathbf{F}$ $\longrightarrow$ (boul-**et**)$_\mathbf{M}$
         'ball' $\longrightarrow$ 'cannonball'

     b. boule$_\mathbf{F}$ $\longrightarrow$ (boul-**ette**)$_\mathbf{F}$
         'ball' $\longrightarrow$ 'small ball'

- The pairs in (2) and (3a) would be unexpected if the suffix simply agreed in gender with the base: rather, it seems that in at least certain cases, the suffix introduces its own gender (a phenomenon documented in other languages, cf. Kramer, 2015).

A more intricate picture (Milner, 1989)

- Milner (1989) however observed that **-ette** may attach to **M**-bases and **-et** to **F**-bases – a phenomenon we dub **gender-mismatch** – leading to a looser semantic relationship between the base and the derived form.

(2) $\text{char}_\mathbf{M}$ $\longrightarrow$ (char-**ette**)$_\mathbf{F}$ (3)   a. $\text{boule}_\mathbf{F}$ $\longrightarrow$ (boul-**et**)$_\mathbf{M}$
         $\text{char}_\mathbf{M}$ *$\longrightarrow$ (char-**et**)$_\mathbf{M}$ 'ball' $\longrightarrow$ 'cannonball'
         'chariot' $\longrightarrow$ 'cart'

                                           b. $\text{boule}_\mathbf{F}$ $\longrightarrow$ (boul-**ette**)$_\mathbf{F}$
                                                  'ball' $\longrightarrow$ 'small ball'

- The pairs in (2) and (3a) would be unexpected if the suffix simply agreed in gender with the base: rather, it seems that in at least certain cases, the suffix introduces its own gender (a phenomenon documented in other languages, cf. Kramer, 2015).

3

Contribution

- In this work, we bring support to a refinement of Milner's observation *via* a more systematic analysis of the French lexicon.

- We also recast the empirical observations in more formal terms, within the framework of Distributed Morphology.

- More specifically, we argue that frequency differences between:
    1. -et and -ette suffixation;
    2. M-to-F *vs* F-to-M gender-mismatches;
    3. the number of "true" diminutives among the -et and -ette forms (regardless of the presence of a gender-mismatch)

- ...could be explained if we assume that:
    - -ette **is ambiguous** between an allomorph of the (non-purely diminutive) suffix -et and another very productive and purely diminutive suffix -ette;
    - gender-mismatching forms results from a **root-level operation**, unlike most gender-matching ones.

Contribution

- In this work, we bring support to a refinement of Milner's observation *via* a more systematic analysis of the French lexicon.

- We also recast the empirical observations in more formal terms, within the framework of Distributed Morphology.

- More specifically, we argue that frequency differences between:
  1. -et and -ette suffixation;
  2. M-to-F *vs* F-to-M gender-mismatches;
  3. the number of "true" diminutives among the -et and -ette forms (regardless of the presence of a gender-mismatch)

- ...could be explained if we assume that:
  - -ette is ambiguous between an allomorph of the (non-purely diminutive) suffix -et and another very productive and purely diminutive suffix -ette;
  - gender-mismatching forms results from a root-level operation, unlike most gender-matching ones.

Contribution

- In this work, we bring support to a refinement of Milner's observation *via* a more systematic analysis of the French lexicon.

- We also recast the empirical observations in more formal terms, within the framework of Distributed Morphology.

- More specifically, we argue that frequency differences between:
  1. -et and -ette suffixation;
  2. M-to-F *vs* F-to-M gender-mismatches;
  3. the number of "true" diminutives among the -et and -ette forms (regardless of the presence of a gender-mismatch)

- ...could be explained if we assume that:
  - -ette is ambiguous between an allomorph of the (non-purely diminutive) suffix -et and another very productive and purely diminutive suffix -ette;
  - gender-mismatching forms results from a root-level operation, unlike most gender-matching ones.

Contribution

- In this work, we bring support to a refinement of Milner's observation *via* a more systematic analysis of the French lexicon.

- We also recast the empirical observations in more formal terms, within the framework of Distributed Morphology.

- More specifically, we argue that frequency differences between:
  1. **-et** and **-ette** suffixation;
  2. **M**-to-**F** *vs* **F**-to-**M** gender-mismatches;
  3. the number of "true" diminutives among the -et and -ette forms (regardless of the presence of a gender-mismatch)

- ...could be explained if we assume that:
  - **-ette is ambiguous** between an allomorph of the (non-purely diminutive) suffix -et and another very productive and purely diminutive suffix -**ette**;
  - gender-mismatching forms results from a **root-level operation**, unlike most gender-matching ones.

Contribution

- In this work, we bring support to a refinement of Milner's observation *via* a more systematic analysis of the French lexicon.

- We also recast the empirical observations in more formal terms, within the framework of Distributed Morphology.

- More specifically, we argue that frequency differences between:
  1. **-et** and **-ette** suffixation;
  2. **M**-to-**F** *vs* **F**-to-**M** gender-mismatches;
  3. the number of "true" diminutives among the -et and -ette forms (regardless of the presence of a gender-mismatch)

- ...could be explained if we assume that:
  - **-ette is ambiguous** between an allomorph of the (non-purely diminutive) suffix -et and another very productive and purely diminutive suffix -ette;
  - gender-mismatching forms results from a **root-level operation**, unlike most gender-matching ones.

Contribution

- In this work, we bring support to a refinement of Milner's observation *via* a more systematic analysis of the French lexicon.
- We also recast the empirical observations in more formal terms, within the framework of Distributed Morphology.
- More specifically, we argue that frequency differences between:
    1. **-et** and **-ette** suffixation;
    2. **M**-to-**F** *vs* **F**-to-**M** gender-mismatches;
    3. the number of "true" diminutives among the **-et** and **-ette** forms (regardless of the presence of a gender-mismatch)
  - ...could be explained if we assume that:
    - **-ette is ambiguous** between an allomorph of the (non-purely diminutive) suffix -et and another very productive and purely diminutive suffix -ette;
    - gender-mismatching forms results from a **root-level operation**, unlike most gender-matching ones.

- In this work, we bring support to a refinement of Milner's observation *via* a more systematic analysis of the French lexicon.
- We also recast the empirical observations in more formal terms, within the framework of Distributed Morphology.
- More specifically, we argue that frequency differences between:
  1. **-et** and **-ette** suffixation;
  2. **M**-to-**F** *vs* **F**-to-**M** gender-mismatches;
  3. the number of "true" diminutives among the **-et** and **-ette** forms (regardless of the presence of a gender-mismatch)
- ...could be explained if we assume that:
  - **-ette is ambiguous** between an allomorph of the (non-purely diminutive) suffix **-et** and another very productive and purely diminutive suffix **-ette**;
  - gender-mismatching forms results from a **root-level operation**, unlike most gender-matching ones.

Contribution

- In this work, we bring support to a refinement of Milner's observation *via* a more systematic analysis of the French lexicon.
- We also recast the empirical observations in more formal terms, within the framework of Distributed Morphology.
- More specifically, we argue that frequency differences between:
    1. **-et** and **-ette** suffixation;
    2. **M**-to-**F** *vs* **F**-to-**M** gender-mismatches;
    3. the number of "true" diminutives among the **-et** and **-ette** forms (regardless of the presence of a gender-mismatch)
- ...could be explained if we assume that:
    - **-ette is ambiguous** between an allomorph of the (non-purely diminutive) suffix **-et** and another very productive and purely diminutive suffix **-ette**;
    - gender-mismatching forms results from a **root-level operation**, unlike most gender-matching ones.

Contribution

- In this work, we bring support to a refinement of Milner's observation *via* a more systematic analysis of the French lexicon.

- We also recast the empirical observations in more formal terms, within the framework of Distributed Morphology.

- More specifically, we argue that frequency differences between:
  1. **-et** and **-ette** suffixation;
  2. **M**-to-**F** *vs* **F**-to-**M** gender-mismatches;
  3. the number of "true" diminutives among the **-et** and **-ette** forms (regardless of the presence of a gender-mismatch)

- ...could be explained if we assume that:
  - **-ette is ambiguous** between an allomorph of the (non-purely diminutive) suffix **-et** and another very productive and purely diminutive suffix **-ette**;
  - gender-mismatching forms results from a **root-level operation**, unlike most gender-matching ones.

Roadmap

Data analysis

    Methodology

    Gender imbalances

    Diminutive semantics

Formal analysis

    Novel claims

    Background on Distributed Morphology

Data analysis

Data collection

- From a list of French words provided by the Linux OS (346,200 entries), we extracted and filtered nouns ending in **-et** and **-ette**.
- Extraction involved:
  - matching all words ending in **-et** or **-ette**;
  - generating, for each candidate form, an approximation of its base by truncation;
  - checking if the approximate base is close enough to an element of the lexicon – if yes, the pair (base, diminutive) is added to our database.
- Filtering involved:
  - finding the exact base from which the word is derived using online resources (Larousse online dictionary, Wiktionary) and introspection;
  - verifying that the base is a nominal.[1]
- The dataset was supplemented by pairs generated *via* pure introspection (not all of them being documented in dictionaries).
- In total, 262 nouns in **-ette** and 146 nouns in **-et** were gathered.

---

[1] 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.

Data collection

- From a list of French words provided by the Linux OS (346,200 entries), we extracted and filtered nouns ending in **-et** and **-ette**.
- Extraction involved:
  - matching all words ending in -et or -ette;
  - generating, for each candidate form, an approximation of its base by truncation;
  - checking if the approximate base is close enough to an element of the lexicon – if yes, the pair (base, diminutive) is added to our database.
- Filtering involved:
  - finding the exact base from which the word is derived using online resources (Larousse online dictionary, Wiktionary) and introspection;
  - verifying that the base is a nominal.[1]
- The dataset was supplemented by pairs generated *via* pure introspection (not all of them being documented in dictionaries).
- In total, 262 nouns in -ette and 146 nouns in -et were gathered.

---

[1]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.

Data collection

- From a list of French words provided by the Linux OS (346,200 entries), we extracted and filtered nouns ending in **-et** and **-ette**.
- Extraction involved:
  - matching all words ending in **-et** or **-ette**;
  - generating, for each candidate form, an approximation of its base by truncation;
  - checking if the approximate base is close enough to an element of the lexicon – if yes, the pair (base, diminutive) is added to our database.
- Filtering involved:
  - finding the exact base from which the word is derived using online resources (Larousse online dictionary, Wiktionary) and introspection;
  - verifying that the base is a nominal.[1]
- The dataset was supplemented by pairs generated *via* pure introspection (not all of them being documented in dictionaries).
- In total, 262 nouns in **-ette** and 146 nouns in **-et** were gathered.

---

[1] 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.

Data collection

- From a list of French words provided by the Linux OS (346,200 entries), we extracted and filtered nouns ending in **-et** and **-ette**.
- Extraction involved:
  - matching all words ending in **-et** or **-ette**;
  - generating, for each candidate form, an approximation of its base by truncation;
  - checking if the approximate base is close enough to an element of the lexicon – if yes, the pair (base, diminutive) is added to our database.
- Filtering involved:
  - finding the exact base from which the word is derived using online resources (Larousse online dictionary, Wiktionary) and introspection;
  - verifying that the base is a nominal.[1]
- The dataset was supplemented by pairs generated *via* pure introspection (not all of them being documented in dictionaries).
- In total, 262 nouns in **-ette** and 146 nouns in **-et** were gathered.

---

[1]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.

Data collection

- From a list of French words provided by the Linux OS (346,200 entries), we extracted and filtered nouns ending in **-et** and **-ette**.
- Extraction involved:
  - matching all words ending in **-et** or **-ette**;
  - generating, for each candidate form, an approximation of its base by truncation;
  - checking if the approximate base is close enough to an element of the lexicon – if yes, the pair (base, diminutive) is added to our database.
- Filtering involved:
  - finding the exact base from which the word is derived using online resources (Larousse online dictionary, Wiktionary) and introspection;
  - verifying that the base is a nominal.[1]
- The dataset was supplemented by pairs generated *via* pure introspection (not all of them being documented in dictionaries).
- In total, 262 nouns in **-ette** and 146 nouns in **-et** were gathered.

---

[1] 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.

Data collection

- From a list of French words provided by the Linux OS (346,200 entries), we extracted and filtered nouns ending in **-et** and **-ette**.
- Extraction involved:
    - matching all words ending in **-et** or **-ette**;
    - generating, for each candidate form, an approximation of its base by truncation;
    - checking if the approximate base is close enough to an element of the lexicon – if yes, the pair (base, diminutive) is added to our database.
- Filtering involved:
    - finding the exact base from which the word is derived using online resources (Larousse online dictionary, Wiktionary) and introspection;
    - verifying that the base is a nominal.[1]
- The dataset was supplemented by pairs generated *via* pure introspection (not all of them being documented in dictionaries).
- In total, 262 nouns in **-ette** and 146 nouns in **-et** were gathered.

---

[1]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.

Data collection

- From a list of French words provided by the Linux OS (346,200 entries), we extracted and filtered nouns ending in **-et** and **-ette**.
- Extraction involved:
  - matching all words ending in **-et** or **-ette**;
  - generating, for each candidate form, an approximation of its base by truncation;
  - checking if the approximate base is close enough to an element of the lexicon – if yes, the pair (base, diminutive) is added to our database.
- Filtering involved:
  - finding the exact base from which the word is derived using online resources (Larousse online dictionary, Wiktionary) and introspection;
  - verifying that the base is a nominal.[1]
- The dataset was supplemented by pairs generated *via* pure introspection (not all of them being documented in dictionaries).
- In total, 262 nouns in **-ette** and 146 nouns in **-et** were gathered.

---

[1]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.

Data collection

- From a list of French words provided by the Linux OS (346,200 entries), we extracted and filtered nouns ending in **-et** and **-ette**.
- Extraction involved:
  - matching all words ending in **-et** or **-ette**;
  - generating, for each candidate form, an approximation of its base by truncation;
  - checking if the approximate base is close enough to an element of the lexicon – if yes, the pair (base, diminutive) is added to our database.
- Filtering involved:
  - finding the exact base from which the word is derived using online resources (Larousse online dictionary, Wiktionary) and introspection;
  - verifying that the base is a nominal.[1]
- The dataset was supplemented by pairs generated *via* pure introspection (not all of them being documented in dictionaries).
- In total, 262 nouns in **-ette** and 146 nouns in **-et** were gathered.

---

[1]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.

Data collection

- From a list of French words provided by the Linux OS (346,200 entries), we extracted and filtered nouns ending in **-et** and **-ette**.
- Extraction involved:
    - matching all words ending in **-et** or **-ette**;
    - generating, for each candidate form, an approximation of its base by truncation;
    - checking if the approximate base is close enough to an element of the lexicon – if yes, the pair (base, diminutive) is added to our database.
- Filtering involved:
    - finding the exact base from which the word is derived using online resources (Larousse online dictionary, Wiktionary) and introspection;
    - verifying that the base is a nominal.[1]
- The dataset was supplemented by pairs generated *via* pure introspection (not all of them being documented in dictionaries).
- In total, 262 nouns in **-ette** and 146 nouns in **-et** were gathered.

---

[1]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.

Data collection

- From a list of French words provided by the Linux OS (346,200 entries), we extracted and filtered nouns ending in **-et** and **-ette**.
- Extraction involved:
  - matching all words ending in **-et** or **-ette**;
  - generating, for each candidate form, an approximation of its base by truncation;
  - checking if the approximate base is close enough to an element of the lexicon – if yes, the pair (base, diminutive) is added to our database.
- Filtering involved:
  - finding the exact base from which the word is derived using online resources (Larousse online dictionary, Wiktionary) and introspection;
  - verifying that the base is a nominal.[1]
- The dataset was supplemented by pairs generated *via* pure introspection (not all of them being documented in dictionaries).
- In total, 262 nouns in **-ette** and 146 nouns in **-et** were gathered.

---

[1]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.

Derived → Base ↓	-ette only	-et only	Both	Total
Feminine	186 (138)	15 (5)	32 (23/7)	233
Masculine	34 (12)	89 (54)	10 (3/6)	133
Total	**220**	**104**	**42**	366

Table 1: Dataset statistics; ▉=gender-preserving suffixation; ▉=**F**-to-**M** mismatches; ▉=**M**-to-**F** mismatches. The number of true diminutives for each count is put in parentheses; for bases with both a -ette and a -et form, the format is (# true -ette diminutives/ # true -et diminutives).

- Quantitatively (and regardless of the gender of the base), **-ette** suffixation is around 1.8 times more frequent than **-et** suffixation ($^{220+42}/_{104+42} \sim 1.8$).
- Qualitatively, generating -ette-forms by introspection appeared easier.
- This all suggests that -ette is overall more productive than -et.

Derived → Base ↓	-ette only	-et only	Both	Total
Feminine	186 (138)	15 (5)	32 (23/7)	233
Masculine	34 (12)	89 (54)	10 (3/6)	133
Total	**220**	**104**	**42**	366

Table 1: Dataset statistics; ▉=gender-preserving suffixation; ▉=**F**-to-**M** mismatches; ▉=**M**-to-**F** mismatches. The number of true diminutives for each count is put in parentheses; for bases with both a -ette and a -et form, the format is (# true -ette diminutives/ # true -et diminutives).

- Quantitatively (and regardless of the gender of the base), **-ette** suffixation is around 1.8 times more frequent than **-et** suffixation ($\frac{220+42}{104+42} \sim 1.8$).
- Qualitatively, generating **-ette**-forms by introspection appeared easier.
- This all suggests that -ette is overall more productive than -et.

Observation 1: -ette vs. -et suffixation

Derived → Base ↓	-ette only	-et only	Both	Total
Feminine	186 (138)	15 (5)	32 (23/7)	233
Masculine	34 (12)	89 (54)	10 (3/6)	133
Total	**220**	**104**	**42**	366

Table 1: Dataset statistics; ▉=gender-preserving suffixation; ▉=**F**-to-**M** mismatches; ▉=**M**-to-**F** mismatches. The number of true diminutives for each count is put in parentheses; for bases with both a -ette and a -et form, the format is (# true -ette diminutives/ # true -et diminutives).

- Quantitatively (and regardless of the gender of the base), **-ette** suffixation is around 1.8 times more frequent than **-et** suffixation ($\frac{220+42}{104+42} \sim 1.8$).
- Qualitatively, generating **-ette**-forms by introspection appeared easier.
- This all suggests that **-ette** is overall more productive than **-et**.

Derived → Base ↓	-ette only	-et only	Both	Total
Feminine	186 (138)	15 (5)	32 (23/7)	233
Masculine	34 (12)	89 (54)	10 (3/6)	133
Total	220	104	42	366

- The proportion of gender-mismatches is higher for M-bases (M-to-F mismatch) than F-bases (F-to-M mismatch):

$$\hat{\mathbb{P}}[\text{-et-form}|\text{F-base}] = {}^{15+32}/{}_{233} = {}^{47}/{}_{233} \sim 20\%$$

$$<_{p=.006}$$

$$\hat{\mathbb{P}}[\text{-ette-form}|\text{M-base}] = {}^{34+10}/{}_{133} = {}^{44}/{}_{133} \sim 33\%$$

- The amplitude of this discrepancy is approximately the same as the one recorded for -et/-ette forms in general (33/20∼1.8).

Derived → Base ↓	-ette only	-et only	Both	Total
Feminine	186 (138)	15 (5)	32 (23/7)	233
Masculine	34 (12)	89 (54)	10 (3/6)	133
Total	220	104	42	366

- The proportion of gender-mismatches is higher for M-bases (M-to-F mismatch) than F-bases (F-to-M mismatch):

$$\hat{\mathbb{P}}[\text{-et-form}|\text{F-base}] = {}^{15+32}/_{233} = {}^{47}/_{233} \sim 20\%$$

$$<_{p=.006}$$

$$\hat{\mathbb{P}}[\text{-ette-form}|\text{M-base}] = {}^{34+10}/_{133} = {}^{44}/_{133} \sim 33\%$$

- The amplitude of this discrepancy is approximately the same as the one recorded for -et/-ette forms in general (33/20∼1.8).

Observation 3: true diminutives × gender-(mis)matches

Derived → Base ↓	-ette only	-et only	Both	Total
Feminine	186 (**138**)	15 (**5**)	32 (**23/7**)	233
Masculine	34 (**12**)	89 (**54**)	10 (**3/6**)	133
Total	220	104	42	366

- More gender-matching forms than gender-mismatching ones appear to have a true diminutive semantics:

$$\hat{\mathbb{P}}[\text{DIM}|\textbf{F}\text{-base-ette or }\textbf{M}\text{-base-et}] = {}^{138+23+54+6}\!/_{186+32+89+10} \sim 70\%$$

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

$$\hat{\mathbb{P}}[\text{DIM}|\textbf{F}\text{-base-et or }\textbf{M}\text{-base-ette}] = {}^{5+7+12+3}\!/_{15+32+34+10} \sim 30\%$$

- This is in line with Milner's observation about the semantic effects of gender-mismatch.

Derived → Base ↓	-ette only	-et only	Both	Total
Feminine	186 (**138**)	15 (**5**)	32 (**23**/**7**)	233
Masculine	34 (**12**)	89 (**54**)	10 (**3**/**6**)	133
Total	220	104	42	366

- More gender-matching forms than gender-mismatching ones appear to have a true diminutive semantics:

$$\hat{\mathbb{P}}[\text{DIM}|\textbf{F}\text{-base-ette or }\textbf{M}\text{-base-et}] = {}^{138+23+54+6}/_{186+32+89+10} \sim 70\%$$

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

$$\hat{\mathbb{P}}[\text{DIM}|\textbf{F}\text{-base-et or }\textbf{M}\text{-base-ette}] = {}^{5+7+12+3}/_{15+32+34+10} \sim 30\%$$

- This is in line with Milner's observation about the semantic effects of gender-mismatch.

Observation 4: a suffix-driven asymmetry

Derived → Base ↓	-ette only	-et only	Both	Total
Feminine	186 (**138**)	15 (5)	32 (**23**/7)	233
Masculine	34 (12)	89 (**54**)	10 (3/**6**)	133
Total	220	104	42	366

- However, an asymmetry driven by the gender of the suffix arises in both "match" and "mismatch" cases!
  - non-mismatching **F**-forms in -ette are more likely to be diminutive than non-mismatching forms in -et:

$$\hat{\mathbb{P}}[\text{DIM}|\textbf{F}\text{-base-ette}] = {}^{138+23}/_{186+32} \sim 74\%$$

$$>_{p=.02}$$

$$\hat{\mathbb{P}}[\text{DIM}|\textbf{M}\text{-base-et}] = {}^{54+6}/_{89+10} \sim 60\%$$

  - the same patterns holds for mismatching forms (although non-significant due to small sample sizes).[2]

---

[2]$\hat{\mathbb{P}}[\text{DIM}|\textbf{M}\text{-base-ette}] = {}^{12+3}/_{34+10} \sim 34\% > \hat{\mathbb{P}}[\text{DIM}|\textbf{F}\text{-base-et}] = {}^{5+7}/_{15+32} \sim 26\%$

Observation 4: a suffix-driven asymmetry

Derived → Base ↓	-ette only	-et only	Both	Total
Feminine	186 (**138**)	15 (5)	32 (**23**/7)	233
Masculine	34 (12)	89 (**54**)	10 (3/**6**)	133
Total	220	104	42	366

- However, an asymmetry driven by the gender of the suffix arises in both "match" and "mismatch" cases!
  - non-mismatching **F**-forms in **-ette** are more likely to be diminutive than non-mismatching forms in **-et**:

$$\hat{\mathbb{P}}[\text{DIM}|\text{F-base-ette}] = {}^{138+23}/_{186+32} \sim 74\%$$

$$>_{p=.02}$$

$$\hat{\mathbb{P}}[\text{DIM}|\text{M-base-et}] = {}^{54+6}/_{89+10} \sim 60\%$$

  - the same patterns holds for mismatching forms (although non-significant due to small sample sizes).[2]

---

[2] $\hat{\mathbb{P}}[\text{DIM}|\text{M-base-ette}] = {}^{12+3}/_{34+10} \sim 34\% > \hat{\mathbb{P}}[\text{DIM}|\text{F-base-et}] = {}^{5+7}/_{15+32} \sim 26\%$

10

Derived → Base ↓	-ette only	-et only	Both	Total
Feminine	186 (**138**)	15 (5)	32 (**23**/7)	233
Masculine	34 (12)	89 (**54**)	10 (3/**6**)	133
Total	220	104	42	366

- However, an asymmetry driven by the gender of the suffix arises in both "match" and "mismatch" cases!
  - non-mismatching **F**-forms in **-ette** are more likely to be diminutive than non-mismatching forms in **-et**:

$$\hat{\mathbb{P}}[\text{DIM}|\textbf{F}\text{-base-}\textbf{ette}] = {}^{138+23}/_{186+32} \sim 74\%$$

$$>_{p=.02}$$

$$\hat{\mathbb{P}}[\text{DIM}|\textbf{M}\text{-base-}\textbf{et}] = {}^{54+6}/_{89+10} \sim 60\%$$

  - the same patterns holds for mismatching forms (although non-significant due to small sample sizes).[2]

---

[2]$\hat{\mathbb{P}}[\text{DIM}|\textbf{M}\text{-base-}\textbf{ette}] = {}^{12+3}/_{34+10} \sim 34\% > \hat{\mathbb{P}}[\text{DIM}|\textbf{F}\text{-base-}\textbf{et}] = {}^{5+7}/_{15+32} \sim 26\%$

Zooming on observation 4: puzzling gender-matching examples

- Even if the gender-matching forms are more likely than the mismatching ones to exhibit a true diminutive semantics, pairs like those in (4) and (5) exemplify some kind of loose semantic relationship, *for both genders*.

(4)   a.   oeil$_\mathbf{M}$ $\rightarrow$ (oeill-*et*)$_\mathbf{M}$
           'eye' $\rightarrow$ 'eyelet'

       b.   arc$_\mathbf{M}$ $\rightarrow$ (arch-*et*)$_\mathbf{M}$
           'bow (archery)' $\rightarrow$ 'bow (music)'

(5)   a.   barre$_\mathbf{F}$ $\rightarrow$ (barr-*ette*)$_\mathbf{F}$
           'bar (construction)' $\rightarrow$ 'hair-clip'

       b.   coquille$_\mathbf{F}$ $\rightarrow$ (coquill-*ette*)$_\mathbf{F}$
           'shell' $\rightarrow$ 'elbow pasta'

- These data suggest 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.

Upshot

- **Effect of gender agreement between the base of the derived form**: gender-matching forms are more likely than gender-mismatching ones to lead to a true diminutive semantics, in line with previous work.

- **Effect of the suffix's gender alone:** -**ette** appears more productive than -**et** and also more likely to lead to a diminutive semantics, and interestingly those two facts somewhat extend to mismatching forms (which were previously thought to be plain lexicalizations).

- We take this last point as evidence that -**ette** is (sometimes, at least) distinct from the allomorph of -**et**.

Formal analysis

- *Contra* previous accounts, we claim that **-ette** is ambiguous between an allomorph of **-et** and a separate (feminine) suffix **-ette**, which we assume is the pure French diminutive suffix DIM, indicating relative smallness, cuteness, or affection towards the object.

- We take that **-et** has a looser semantics, which only involves a similarity with the base w.r.t. a salient feature, usually shape (so we write **-et** = SHAPE for brevity). Milner (1989) and Delhay (1999) had made a similar observation already, but mostly for cases of gender-mismatch.

$$\begin{array}{ccc}
 & \textbf{M} \\
\textsc{Shape} & \nearrow & \text{-et} \\
 & \textbf{F} \searrow \\
\textsc{Dim} & \longrightarrow & \text{-ette}
\end{array}$$

Novel claims about the diminutive suffixes -et and -ette

- *Contra* previous accounts, we claim that **-ette** is ambiguous between an allomorph of **-et** and a separate (feminine) suffix **-ette**, which we assume is the pure French diminutive suffix DIM, indicating relative smallness, cuteness, or affection towards the object.

- We take that **-et** has a looser semantics, which only involves a similarity with the base w.r.t. a salient feature, usually shape (so we write **-et** = SHAPE for brevity). Milner (1989) and Delhay (1999) had made a similar observation already, but mostly for cases of gender-mismatch.

SHAPE   **M** → **-et**

      F

DIM ⟶ **-ette**

Novel claims about the diminutive suffixes -et and -ette

- *Contra* previous accounts, we claim that **-ette** is ambiguous between an allomorph of **-et** and a separate (feminine) suffix **-ette**, which we assume is the pure French diminutive suffix DIM, indicating relative smallness, cuteness, or affection towards the object.

- We take that **-et** has a looser semantics, which only involves a similarity with the base w.r.t. a salient feature, usually shape (so we write **-et** = SHAPE for brevity). Milner (1989) and Delhay (1999) had made a similar observation already, but mostly for cases of gender-mismatch.

- **-et** being the realization of Shape and **-ette** being that of either Shape+Agree or Dim can explain why **-ette** is more frequent than **-et** across the board, and more likely to yield a diminutive semantics.

- Now we have to devise the exact conditions under which each exponent can surface, to account for the fact that gender mismatching forms overall yield less true diminutives. To do so, let's take a detour to Distributed Morphology.

- **-et** being the realization of SHAPE and **-ette** being that of either SHAPE+AGREE or DIM can explain why **-ette** is more frequent than **-et** across the board, and more likely to yield a diminutive semantics.

- Now we have to devise the exact conditions under which each exponent can surface, to account for the fact that gender mismatching forms overall yield less true diminutives. To do so, let's take a detour to Distributed Morphology.

Background on Distributed Morphology

- Distributed Morphology (**DM**, Halle and Marantz (1993)) assumes that there is no divide between syntax and morphology.
- The functions that other theories ascribe to the "lexicon" are "distributed" among other components of the grammar:
  - the **Formative List** provides the input for syntax in the form of feature bundles (no phonology at that point!);
  - **syntactic operations** (MERGE, MOVE, AGREE) apply to formatives;
  - the **Exponent List** is used for (post-syntactic) vocabulary insertion.
- In DM, **roots (formatives) are proposed to be category-neutral**. They undergo categorization by functional elements or *heads*: *n*, *a*, *v*...

(a) The noun 'hammer'   (b) The verb 'to hammer'

Figure 1: Categorizing a root in different ways in DM

Background on Distributed Morphology

- Distributed Morphology (**DM**, Halle and Marantz (1993)) assumes that there is no divide between syntax and morphology.
- The functions that other theories ascribe to the "lexicon" are "distributed" among other components of the grammar:
  - the **Formative List** provides the input for syntax in the form of feature bundles (no phonology at that point!);
  - **syntactic operations** (MERGE, MOVE, AGREE) apply to formatives;
  - the **Exponent List** is used for (post-syntactic) vocabulary insertion.
- In DM, **roots (formatives) are proposed to be category-neutral**. They undergo categorization by functional elements or *heads*: *n*, *a*, *v*...

(a) The noun 'hammer' (b) The verb 'to hammer'

Figure 1: Categorizing a root in different ways in DM

Background on Distributed Morphology

- Distributed Morphology (**DM**, Halle and Marantz (1993)) assumes that there is no divide between syntax and morphology.
- The functions that other theories ascribe to the "lexicon" are "distributed" among other components of the grammar:
    - the **Formative List** provides the input for syntax in the form of feature bundles (no phonology at that point!);
    - **syntactic operations** (MERGE, MOVE, AGREE) apply to formatives;
    - the **Exponent List** is used for (post-syntactic) vocabulary insertion.
- In DM, **roots (formatives) are proposed to be category-neutral**. They undergo categorization by functional elements or *heads*: *n*, *a*, *v*...

*n*p

√hammer   *n*

(a) The noun 'hammer'

*v*p

√hammer   *v*

(b) The verb 'to hammer'

Figure 1: Categorizing a root in different ways in DM

Background on Distributed Morphology

- Distributed Morphology (**DM**, Halle and Marantz (1993)) assumes that there is no divide between syntax and morphology.
- The functions that other theories ascribe to the "lexicon" are "distributed" among other components of the grammar:
    - the **Formative List** provides the input for syntax in the form of feature bundles (no phonology at that point!);
    - **syntactic operations** (MERGE, MOVE, AGREE) apply to formatives;
    - the **Exponent List** is used for (post-syntactic) vocabulary insertion.
- In DM, **roots (formatives) are proposed to be category-neutral**. They undergo categorization by functional elements or *heads*: *n*, *a*, *v*...
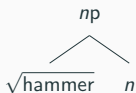

(a) The noun 'hammer'   (b) The verb 'to hammer'
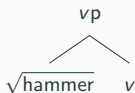
Figure 1: Categorizing a root in different ways in DM

Background on Distributed Morphology

- Distributed Morphology (**DM**, Halle and Marantz (1993)) assumes that there is no divide between syntax and morphology.
- The functions that other theories ascribe to the "lexicon" are "distributed" among other components of the grammar:
    - the **Formative List** provides the input for syntax in the form of feature bundles (no phonology at that point!);
    - **syntactic operations** (MERGE, MOVE, AGREE) apply to formatives;
    - the **Exponent List** is used for (post-syntactic) vocabulary insertion.
- In DM, **roots (formatives) are proposed to be category-neutral**. They undergo categorization by functional elements or *heads*: *n*, *a*, *v*...

(a) The noun 'hammer'

(b) The verb 'to hammer'

Figure 1: Categorizing a root in different ways in DM

Background on Distributed Morphology

- Distributed Morphology (**DM**, Halle and Marantz (1993)) assumes that there is no divide between syntax and morphology.
- The functions that other theories ascribe to the "lexicon" are "distributed" among other components of the grammar:
    - the **Formative List** provides the input for syntax in the form of feature bundles (no phonology at that point!);
    - **syntactic operations** (MERGE, MOVE, AGREE) apply to formatives;
    - the **Exponent List** is used for (post-syntactic) vocabulary insertion.
- In DM, **roots (formatives) are proposed to be category-neutral**. They undergo categorization by functional elements or *heads*: *n*, *a*, *v*...

(a) The noun 'hammer'

(b) The verb 'to hammer'

Figure 1: Categorizing a root in different ways in DM

DM and the semantics-morphology interface

- Subsequent work within the DM framework (Marantz, 1997, 2001; Arad, 2003, 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 distinction can successfully explain **opacity effects witnessed in both the semantic and phonological domain**.

  - For instance, affixes which shift stress (like *-ity*) also tend to lead to semantically less predictable meanings than otherwise similar affixes, (like *-ness*) which do not shift stress (Aronoff, 1976).

  - The former kind of affix merges at the root level (where form and meaning are still accessible and modifiable), while the latter kind merges with already created words, whose phonology and semantics are fixed.

fátal

√fátal   *a*

(a) The adjective 'fátal'

fatálity

√fátal   *-ity*

(b) The noun 'fatálity'

fátalness

          -ness

√fátal   *a*

(c) The noun 'fátalness'

DM and the semantics-morphology interface

- Subsequent work within the DM framework (Marantz, 1997, 2001; Arad, 2003, 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 distinction can successfully explain **opacity effects witnessed in both the semantic and phonological domain**.

  - For instance, affixes which shift stress (like *-ity*) also tend to lead to semantically less predictable meanings than otherwise similar affixes, (like *-ness*) which do not shift stress (Aronoff, 1976).

  - The former kind of affix merges at the root level (where form and meaning are still accessible and modifiable), while the latter kind merges with already created words, whose phonology and semantics are fixed.

fátal

√fátal   *a*

(a) The adjective 'fátal'

fatálity

√fátal   *-ity*

(b) The noun 'fatálity'

fátalness

        -ness

√fátal   *a*

(c) The noun 'fátalness'

16

DM and the semantics-morphology interface

- Subsequent work within the DM framework (Marantz, 1997, 2001; Arad, 2003, 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 distinction can successfully explain **opacity effects witnessed in both the semantic and phonological domain**.

  - For instance, affixes which shift stress (like *-ity*) also tend to lead to semantically less predictable meanings than otherwise similar affixes, (like *-ness*) which do not shift stress (Aronoff, 1976).

  - The former kind of affix merges at the root level (where form and meaning are still accessible and modifiable), while the latter kind merges with already created words, whose phonology and semantics are fixed.

fátal

√fátal   a

(a) The adjective 'fátal'

fatálity

√fátal   -ity

(b) The noun 'fatálity'

fátalness

          -ness

√fátal   a

(c) The noun 'fátalness'

16

DM and the semantics-morphology interface

- Subsequent work within the DM framework (Marantz, 1997, 2001; Arad, 2003, 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 distinction can successfully explain **opacity effects witnessed in both the semantic and phonological domain**.

  - For instance, affixes which shift stress (like *-ity*) also tend to lead to semantically less predictable meanings than otherwise similar affixes, (like *-ness*) which do not shift stress (Aronoff, 1976).

  - The former kind of affix merges at the root level (where form and meaning are still accessible and modifiable), while the latter kind merges with already created words, whose phonology and semantics are fixed.

fátal

√fátal   *a*

(a) The adjective 'fátal'

fatálity

√fátal   *-ity*

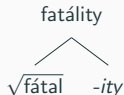(b) The noun 'fatálity'

fátalness

    -ness

√fátal   *a*

(c) The noun 'fátalness'

16

DM and the semantics-morphology interface

- Subsequent work within the DM framework (Marantz, 1997, 2001; Arad, 2003, 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 distinction can successfully explain **opacity effects witnessed in both the semantic and phonological domain**.

  - For instance, affixes which shift stress (like *-ity*) also tend to lead to semantically less predictable meanings than otherwise similar affixes, (like *-ness*) which do not shift stress (Aronoff, 1976).

  - The former kind of affix merges at the root level (where form and meaning are still accessible and modifiable), while the latter kind merges with already created words, whose phonology and semantics are fixed.

fátal

(a) The adjective 'fátal'
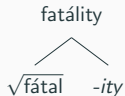
fatálity

(b) The noun 'fatálity'

fátalness

(c) The noun 'fátalness'

DM and the semantics-morphology interface

- Subsequent work within the DM framework (Marantz, 1997, 2001; Arad, 2003, 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 distinction can successfully explain **opacity effects witnessed in both the semantic and phonological domain**.

  - For instance, affixes which shift stress (like *-ity*) also tend to lead to semantically less predictable meanings than otherwise similar affixes, (like *-ness*) which do not shift stress (Aronoff, 1976).

  - The former kind of affix merges at the root level (where form and meaning are still accessible and modifiable), while the latter kind merges with already created words, whose phonology and semantics are fixed.
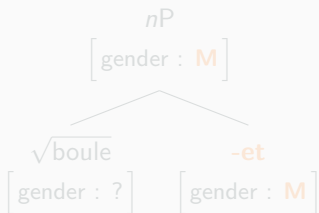
fátal

(a) The adjective 'fátal'

fatálity

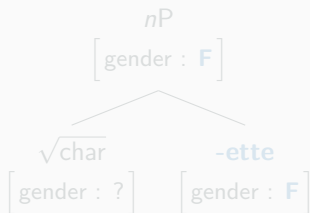**(b)** The noun 'fatálity'

fátalness

(c) The noun 'fátalness'

16

DM and the semantics-morphology interface

- Subsequent work within the DM framework (Marantz, 1997, 2001; Arad, 2003, 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 distinction can successfully explain **opacity effects witnessed in both the semantic and phonological domain**.

  - For instance, affixes which shift stress (like *-ity*) also tend to lead to semantically less predictable meanings than otherwise similar affixes, (like *-ness*) which do not shift stress (Aronoff, 1976).

  - The former kind of affix merges at the root level (where form and meaning are still accessible and modifiable), while the latter kind merges with already created words, whose phonology and semantics are fixed.

fátal

$\sqrt{\text{fátal}}$ a

(a) The adjective 'fátal'

fatálity

$\sqrt{\text{fátal}}$ *-ity*

(b) The noun 'fatálity'

fátalness

-ness

$\sqrt{\text{fátal}}$ a

(c) The noun 'fátalness'

16

Accounting for gender-mismatching forms

- Following this literature, we want to suggest that gender-mismatching forms result from a merger of the DIM (exponent:-**ette**) or SHAPE (exponent:-**et**) suffix with mismatching gender at the root-level.

- At that level, the suffix plays the role of a categorizing head and can superimpose its own gender to the category-neutral root.

- Also, both suffixes (but SHAPE in particular), act on semantically underspecified roots, and as such lead to looser meanings.
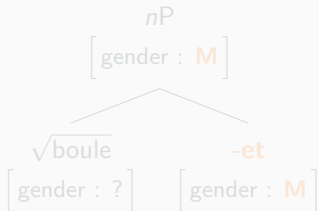
(a) Something akin to the root $\sqrt{\text{ball}}$ ($\simeq$ ball) in terms of SHAPE
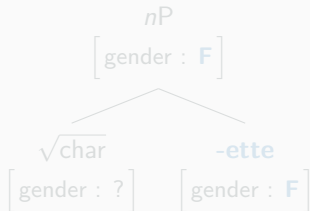
(b) A diminutive of the root $\sqrt{\text{chariot}}$ ($\simeq$ chariot)

Accounting for gender-mismatching forms

- Following this literature, we want to suggest that gender-mismatching forms result from a merger of the DIM (exponent:-**ette**) or SHAPE (exponent:-**et**) suffix with mismatching gender at the root-level.
- At that level, the suffix plays the role of a categorizing head and can superimpose its own gender to the category-neutral root.
- Also, both suffixes (but SHAPE in particular), act on semantically underspecified roots, and as such lead to looser meanings.

nP
$\left[\text{gender} : \textbf{M} \right]$

$\sqrt{\text{boule}}$ -**et**
$\left[\text{gender} : \text{?} \right]$ $\left[\text{gender} : \textbf{M} \right]$

(a) Something akin to the root $\sqrt{\text{ball}}$ ($\simeq$ ball) in terms of SHAPE

nP
$\left[\text{gender} : \textbf{F} \right]$

$\sqrt{\text{char}}$ -**ette**
$\left[\text{gender} : \text{?} \right]$ $\left[\text{gender} : \textbf{F} \right]$

(b) A diminutive of the root $\sqrt{\text{chariot}}$ ($\simeq$ chariot)

17

Accounting for gender-mismatching forms

- Following this literature, we want to suggest that gender-mismatching forms result from a merger of the DIM (exponent:**-ette**) or SHAPE (exponent:**-et**) suffix with mismatching gender at the root-level.
- At that level, the suffix plays the role of a categorizing head and can superimpose its own gender to the category-neutral root.
- Also, both suffixes (but SHAPE in particular), act on semantically underspecified roots, and as such lead to looser meanings.

(a) Something akin to the root √ball (≃ ball) in terms of SHAPE

(b) A diminutive of the root √chariot (≃ chariot)

17

Accounting for gender-matching forms

- Gender-matching forms have their suffix merged above the nominalizing head *n* (which we assume hosts gender features). In that case the suffix simply has to agree in gender with the already categorized noun.

- Problem: some matching forms could very well result from a root-level derivation whereby the gender of the categorizing suffix incidentally matches that of the noun (if it were categorized)…

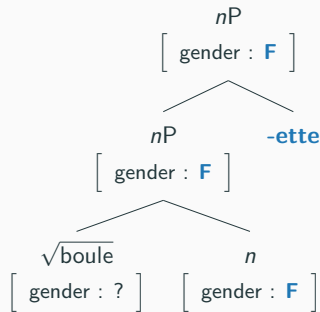- We want to block that kind of configuration by saying that affixes should be merged as high as possible: "Merge high!".

Figure 4: A diminutive of the categorized noun 'ball' (feminine).

Accounting for gender-matching forms

- Gender-matching forms have their suffix merged above the nominalizing head n (which we assume hosts gender features). In that case the suffix simply has to agree in gender with the already categorized noun.

- Problem: some matching forms could very well result from a root-level derivation whereby the gender of the categorizing suffix incidentally matches that of the noun (if it were categorized)...

- We want to block that kind of configuration by saying that affixes should be merged as high as possible: "Merge high!".
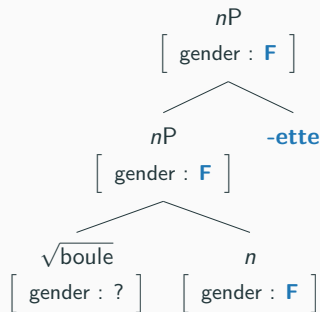

Figure 4: A diminutive of the categorized noun 'ball' (feminine).

Accounting for gender-matching forms

- Gender-matching forms have their suffix merged above the nominalizing head n (which we assume hosts gender features). In that case the suffix simply has to agree in gender with the already categorized noun.

- Problem: some matching forms could very well result from a root-level derivation whereby the gender of the categorizing suffix incidentally matches that of the noun (if it were categorized)...

- **We want to block that kind of configuration by saying that affixes should be merged as high as possible: "Merge high!"**.
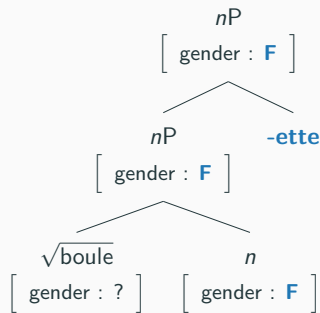
Figure 4: A diminutive of the categorized noun 'ball' (feminine).

Further predictions

- This analysis, making use of both exponent ambiguity and of the existence of two distinct levels of morphological derivation, explains why gender-mismatching forms are less likely to be diminutive, *while still exhibiting a gender-related asymmetry*.

- In particular, we predict **M**-to-**F** forms in -**ette** to exhibit a diminutive semantics (contributed by -**ette**, which is unambiguously DIM in that case), but not on the "right"/most salient entity (due to root-underspecification). This might be the case for the pairs below.

(6) a. cigare$_\mathbf{M}$ → (cigar-***ette***)$_\mathbf{F}$
         'cigar' → 'cigarette'

     b. disque$_\mathbf{M}$ → (disqu-***ette***)$_\mathbf{F}$
         'CD, (hard) disk' → 'floppy disk'

Summary

base	suffix	level	form	semantics
M	SHAPE	~~1~~/2	**-et**	loose on "exact" base
	DIM	1	**-ette**	diminutive on "noisy" root
F	SHAPE	1	**-et**	loose on "noisy" root
	SHAPE+AGR	2	**-ette**	loose on "exact" base
	DIM	~~1~~/2		diminutive on "exact" base

Table 2: Summary of the predictions. '1'=root-level derivation; '2'=above n. Struckthrough level are ruled out by our condition "MERGE high!".

Conclusion

- We argued that the difference in productivity and transparency between -**ette** and -**et** was due to -**ette** being **ambiguous between an allomorph of -et (not purely diminutive) and Dim**.

- We showed the discrepancy was modulated by gender-mismatches, which we argued were the result of **root-level derivation and therefore linked to extra semantic noise**.

- Crucially, our account provided a morphosyntactic explanation as to why **gender-mismatches correlate with some form of semantic mismatch**; accounts positing lexicalization did not address this.

- However, it comes at the cost of assuming that **the suffixes Dim and Shape can in principle merge either at the root-level or higher** (modulo the "MERGE high!" condition). This is unwarranted, but not totally unprecedented (cf. *-ity* affixed unexpectedly "high" after *-able*, Aronoff and Lindsay (2014)).

- Another limit is that the account, in predicting the gender imbalance in the dataset, may **overgenerate in the semantic domain** (cf. previous Table). This is hard to assess as judgments are subtle.

Conclusion

- We argued that the difference in productivity and transparency between -**ette** and -**et** was due to -**ette** being **ambiguous between an allomorph of -et (not purely diminutive) and Dim**.
- We showed the discrepancy was modulated by gender-mismatches, which we argued were the result of **root-level derivation and therefore linked to extra semantic noise**.
- Crucially, our account provided a morphosyntactic explanation as to why **gender-mismatches correlate with some form of semantic mismatch**; accounts positing lexicalization did not address this.
- However, it comes at the cost of assuming that **the suffixes Dim and Shape can in principle merge either at the root-level or higher** (modulo the "MERGE high!" condition). This is unwarranted, but not totally unprecedented (cf. *-ity* affixed unexpectedly "high" after *-able*, Aronoff and Lindsay (2014)).
- Another limit is that the account, in predicting the gender imbalance in the dataset, may **overgenerate in the semantic domain** (cf. previous Table). This is hard to assess as judgments are subtle.

Conclusion

- We argued that the difference in productivity and transparency between -**ette** and -**et** was due to -**ette** being **ambiguous between an allomorph of -et (not purely diminutive) and Dim**.
- We showed the discrepancy was modulated by gender-mismatches, which we argued were the result of **root-level derivation and therefore linked to extra semantic noise**.
- Crucially, our account provided a morphosyntactic explanation as to why **gender-mismatches correlate with some form of semantic mismatch**; accounts positing lexicalization did not address this.
- However, it comes at the cost of assuming that **the suffixes Dim and Shape can in principle merge either at the root-level or higher** (modulo the "MERGE high!" condition). This is unwarranted, but not totally unprecedented (cf. -*ity* affixed unexpectedly "high" after -*able*, Aronoff and Lindsay (2014)).
- Another limit is that the account, in predicting the gender imbalance in the dataset, may **overgenerate in the semantic domain** (cf. previous Table). This is hard to assess as judgments are subtle.

Conclusion

- We argued that the difference in productivity and transparency between -**ette** and -**et** was due to -**ette** being **ambiguous between an allomorph of -et (not purely diminutive) and Dim**.
- We showed the discrepancy was modulated by gender-mismatches, which we argued were the result of **root-level derivation and therefore linked to extra semantic noise**.
- Crucially, our account provided a morphosyntactic explanation as to why **gender-mismatches correlate with some form of semantic mismatch**; accounts positing lexicalization did not address this.
- However, it comes at the cost of assuming that **the suffixes Dim and Shape can in principle merge either at the root-level or higher** (modulo the "MERGE high!" condition). This is unwarranted, but not totally unprecedented (cf. -*ity* affixed unexpectedly "high" after -*able*, Aronoff and Lindsay (2014)).
- Another limit is that the account, in predicting the gender imbalance in the dataset, may **overgenerate in the semantic domain** (cf. previous Table). This is hard to assess as judgments are subtle.

Conclusion

- We argued that the difference in productivity and transparency between -**ette** and -**et** was due to -**ette** being **ambiguous between an allomorph of -et (not purely diminutive) and Dim**.
- We showed the discrepancy was modulated by gender-mismatches, which we argued were the result of **root-level derivation and therefore linked to extra semantic noise**.
- Crucially, our account provided a morphosyntactic explanation as to why **gender-mismatches correlate with some form of semantic mismatch**; accounts positing lexicalization did not address this.
- However, it comes at the cost of assuming that **the suffixes Dim and Shape can in principle merge either at the root-level or higher** (modulo the "MERGE high!" condition). This is unwarranted, but not totally unprecedented (cf. -*ity* affixed unexpectedly "high" after -*able*, Aronoff and Lindsay (2014)).
- Another limit is that the account, in predicting the gender imbalance in the dataset, may **overgenerate in the semantic domain** (cf. previous Table). This is hard to assess as judgments are subtle.

Outlook

- Can/Should this kind of analysis extend to other pairs of French suffixes with gender inflections (e.g. the *-ier*/*-ière* pair mentioned yesterday, or the sometimes augmentative, sometimes diminutive suffix *-ot*)?

- What about other syntactic categories, such as adjectives (e.g. *mignonet*, *longuet*, *gentillet*)?

- We may try to approximate probabilities for each possible parse (with DIM, SHAPE, high or low), to see if they more or less match our empirical statistics.

- Lastly, we could try to see if the semantic relations we talked about are somewhat captured by statistical models of language such as word embedding models, which come with ways to quantitatively measure semantic similarity as well as the variability in the "direction" ($\sim$ meaning) of morphological transformations ***time for shameless self-citations...*** (Benbaji et al., 2022; Hénot-Mortier, 2022; Benbaji et al., 2023).

22

Outlook

- Can/Should this kind of analysis extend to other pairs of French suffixes with gender inflections (e.g. the *-ier*/*-ière* pair mentioned yesterday, or the sometimes augmentative, sometimes diminutive suffix *-ot*)?
- What about other syntactic categories, such as adjectives (e.g. *mignonet*, *longuet*, *gentillet*)?
- We may try to approximate probabilities for each possible parse (with DIM, SHAPE, high or low), to see if they more or less match our empirical statistics.
- Lastly, we could try to see if the semantic relations we talked about are somewhat captured by statistical models of language such as word embedding models, which come with ways to quantitatively measure semantic similarity as well as the variability in the "direction" ($\sim$ meaning) of morphological transformations ***time for shameless self-citations...*** (Benbaji et al., 2022; Hénot-Mortier, 2022; Benbaji et al., 2023).

Outlook

- Can/Should this kind of analysis extend to other pairs of French suffixes with gender inflections (e.g. the *-ier/-ière* pair mentioned yesterday, or the sometimes augmentative, sometimes diminutive suffix *-ot*)?
- What about other syntactic categories, such as adjectives (e.g. *mignonet*, *longuet*, *gentillet*)?
- We may try to approximate probabilities for each possible parse (with DIM, SHAPE, high or low), to see if they more or less match our empirical statistics.
- Lastly, we could try to see if the semantic relations we talked about are somewhat captured by statistical models of language such as word embedding models, which come with ways to quantitatively measure semantic similarity as well as the variability in the "direction" ($\sim$ meaning) of morphological transformations ***time for shameless self-citations...*** (Benbaji et al., 2022; Hénot-Mortier, 2022; Benbaji et al., 2023).

Outlook

- Can/Should this kind of analysis extend to other pairs of French suffixes with gender inflections (e.g. the *-ier/-ière* pair mentioned yesterday, or the sometimes augmentative, sometimes diminutive suffix *-ot*)?
- What about other syntactic categories, such as adjectives (e.g. *mignonet*, *longuet*, *gentillet*)?
- We may try to approximate probabilities for each possible parse (with DIM, SHAPE, high or low), to see if they more or less match our empirical statistics.
- Lastly, we could try to see if the semantic relations we talked about are somewhat captured by statistical models of language such as word embedding models, which come with ways to quantitatively measure semantic similarity as well as the variability in the "direction" ($\sim$ meaning) of morphological transformations ***time for shameless self-citations...*** (Benbaji et al., 2022; Hénot-Mortier, 2022; Benbaji et al., 2023).

Thank you very much for your attention !

Gougenheim, G. (1946). Les feminins diminutifs en francais moderne. Modern Language Notes, 61(6), 416. https://doi.org/10.2307/2908930

Aronoff, M. (1976). Word formation in generative grammar.

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.

Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), The view from building 20 (pp. 111–176). MIT Press.

Jurafsky, D. (1996). Universal tendencies in the semantics of the diminutive. Language, 72(3), 533. https://doi.org/10.2307/416278

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

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

Dressler, W. U., & Barbaresi, L. M. (2001). Morphopragmatics of diminutives and augmentatives. In Perspectives on semantics, pragmatics, and discourse (pp. 43–58). John Benjamins Publishing Company. https://doi.org/10.1075/pbns.90.07dre

Selected references ii

Marantz, A. (2001). Words.

Plénat, M., & Roché, M. (2001). Prosodic constraints on suffixation in French. Proceedings of the Third Mediterranean Morphology Meeting.

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

Schneider, K. P. (2003). M - sn. Max Niemeyer Verlag. https://doi.org/doi:10.1515/9783110929553

Arad, M. (2005). Roots and patterns: Hebrew morpho-syntax. Springer.

Ferrari, F. (2005). A syntactic analysis of the nominal systems of italian and luganda: How nouns can be formed in (Doctoral dissertation). New York University.

Nelson, D. (2005). French gender assignment revisited. Word, 56(1), 19–38.

Lyster, R. (2006). Predictability in french gender attribution: A corpus analysis. Journal of French Language Studies, 16(1), 69–92. https://doi.org/10.1017/S0959269506002304

Booij, G. (2007). The Grammar of Words: An Introduction to Linguistic Morphology. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199226245.001.0001

Kornexl, L. (2008). Women and other 'small things': -ette as a feminine marker. In English historical linguistics 2006 (pp. 241–257). John Benjamins Publishing Company. https://doi.org/10.1075/cilt.296.16kor

Aronoff, M., & Lindsay, M. (2014). 67Productivity, Blocking, and Lexicalization. In The Oxford Handbook of Derivational Morphology. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199641642.013.0005

Kramer, R. (2015). The morphosyntax of gender. Oxford University Press.

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

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

Benbaji, I., Doron, O., & Hénot-Mortier, A. (2022). Word-embeddings distinguish denominal and root-derived verbs in semitic. Proceedings of the Workshop on End-to-End Compositional Models of Vector-Based Semantics.

Hénot-Mortier, A. (2022). Evidence for an encoding of morphological blocking effects within two english word embedding models. Architectures and Mechanisms of Language Processing 2022 (AMLaP 28).

Benbaji, I., Doron, O., & Hénot-Mortier, A. (2023). Distinguishing levels of morphological derivation in word-embedding models.
Proceedings of the 53rd Annual Meeting of the North East Linguistic Society (NELS 53).

A remaining puzzle(?)

- Why are [60]/[99] **M**-forms in **-et** diminutive, given that we predict the more general SHAPE relationship to hold in that case?
- We think this may be due to some form of morphological reanalysis targeting a specific subset of the *-et*-forms.
- Indeed, a DIM-meaning is more likely to arise for bases ending in *in*/*on*/*eau* (38/41), which already have a fossilized diminutive flavor:[3]
- Such endings were also the preferred targets for applying **-et** productively.
- This suggests that they were perhaps re-analyzed as proper morphemes (contributing the DIM semantics) by the action of **-et** suffixation.

---

[3]We use this denomination because most of the nominals from the dataset with such endings (*e.g. cochon*, 'pig', *champignon*, 'mushroom') were morphologically simplex; yet, the same endings are common in proper names (*Antoine → Antonin*; *Marie → Marion*; *Boucher → Bouchereau*...) and appear consistently diminutive.