

Contrasting clippings in English and French: A constructional perspective

Martin Hilpert

David Correia Saavedra

Jennifer Rains

prof

gator

fridge

professor

alligator

refrigerator

aprèm

dwich

soce

après-midi

sandwich

associé

Are there regularities in the way
speakers form clippings?

Are there regularities in the way speakers form clippings?

Can we explain when and why speakers clip different word parts?

Are there regularities in the way speakers form clippings?

Can we explain when and why speakers clip different word parts?

Are there systematic differences between English and French?

the clipping database



2400 English clippings

2100 French clippings

annotations for phonological,
morphological, semantic, and
frequency-based variables

analysis of typical configurations

the clipping database



2400 English clippings
2100 French clippings

annotations for phonological,
morphological, semantic, and
frequency-based variables

analysis of typical configurations

experiments



forced choice
between consonant-final
and vowel-final clippings

production task
with source word as prompt

the clipping database



2400 English clippings
2100 French clippings

annotations for phonological,
morphological, semantic, and
frequency-based variables

analysis of typical configurations

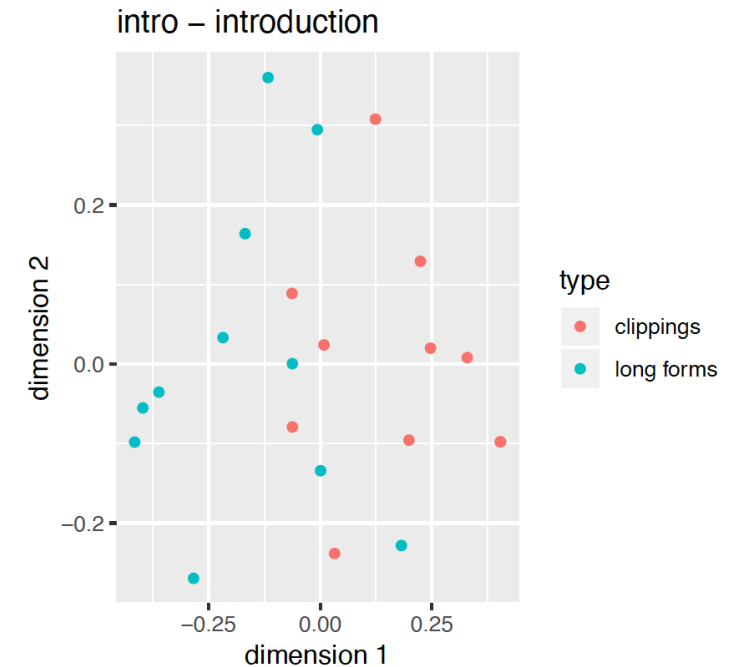
experiments



forced choice
between consonant-final
and vowel-final clippings

production task
with source word as prompt

corpus analyses



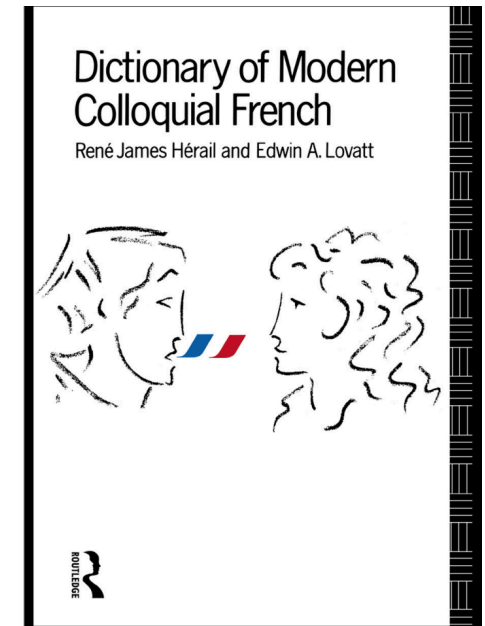
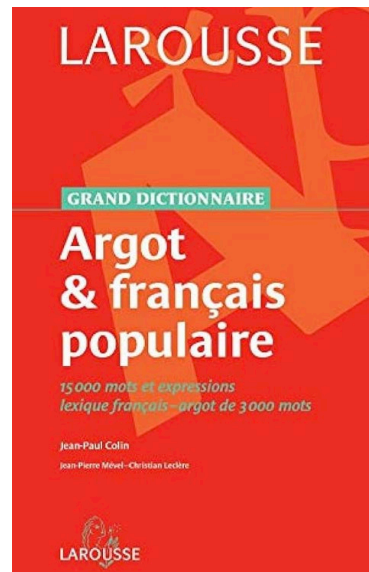
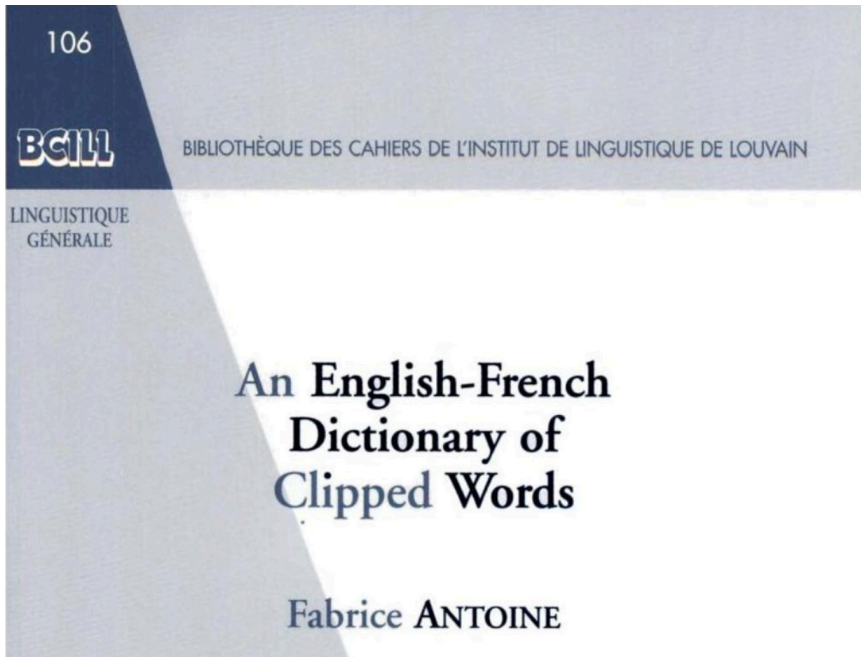
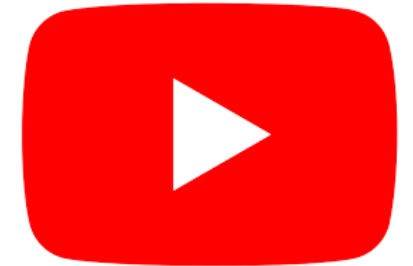
distributional semantics

meaning differences
between clippings
and their source words

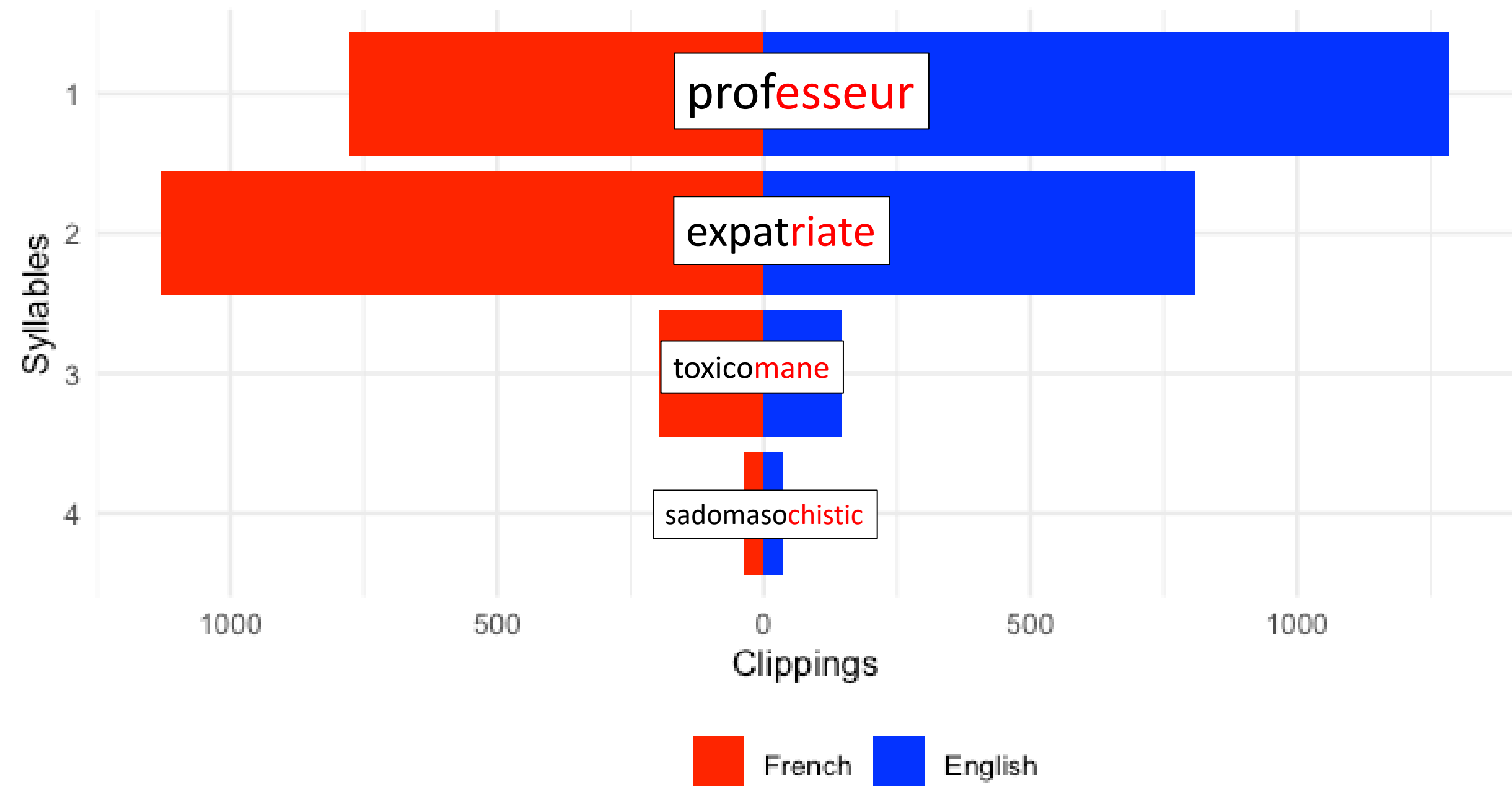
Overview of this talk

1. The clipping database and what we can learn from it
2. Contrasting English and French
3. Discussion and conclusions

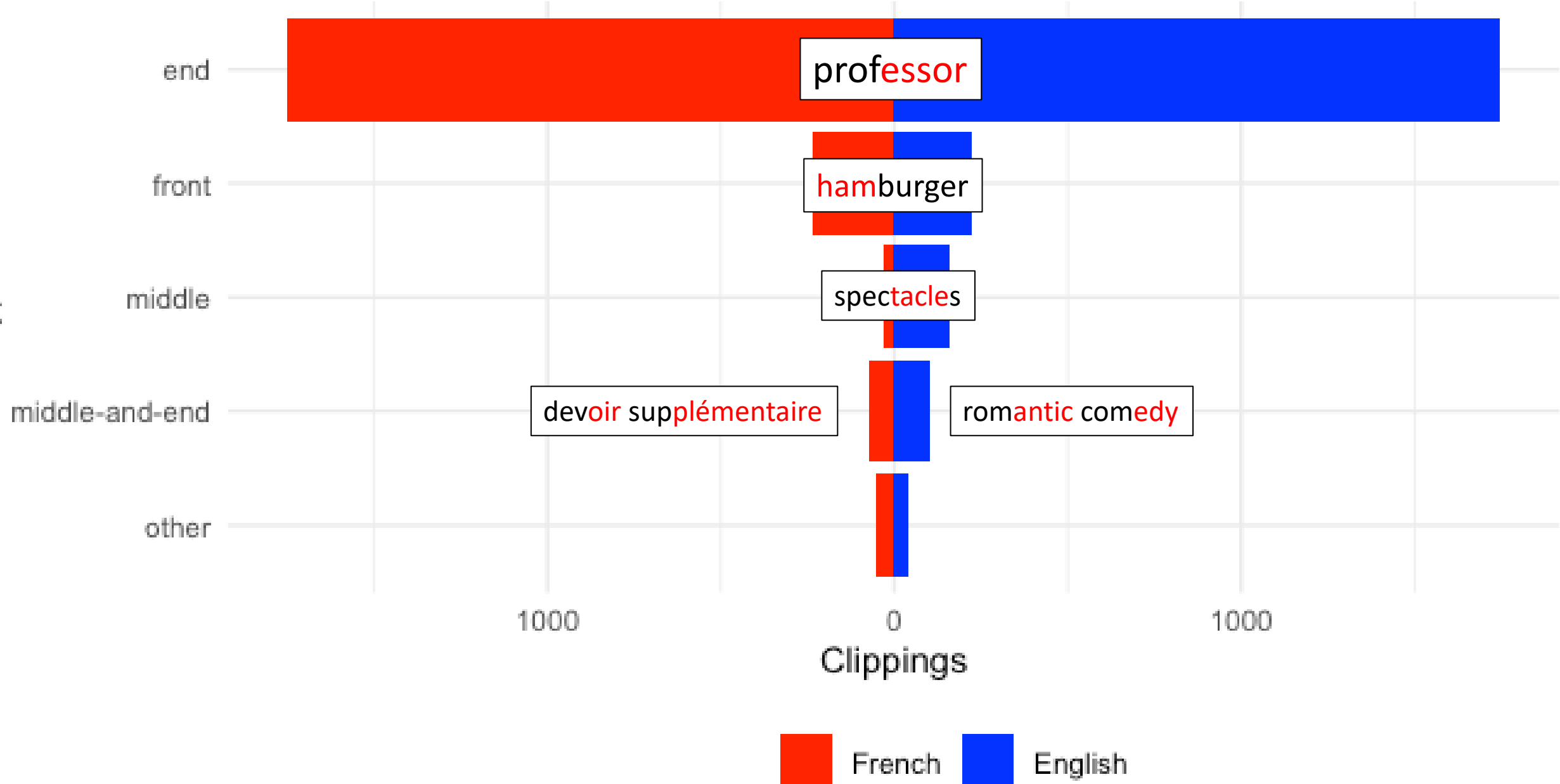
1. The clipping database and what we can learn from it



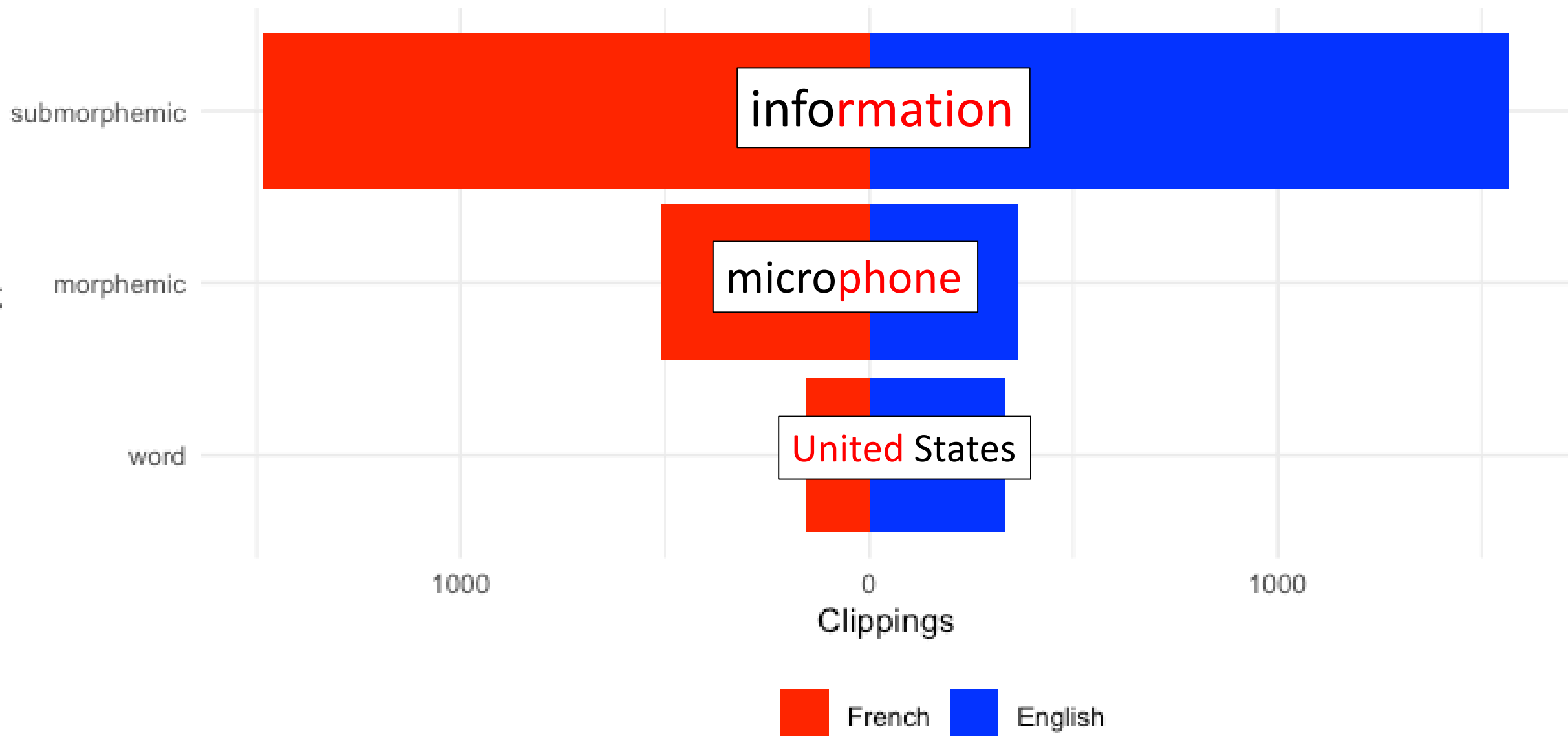
Clipping length in syllables



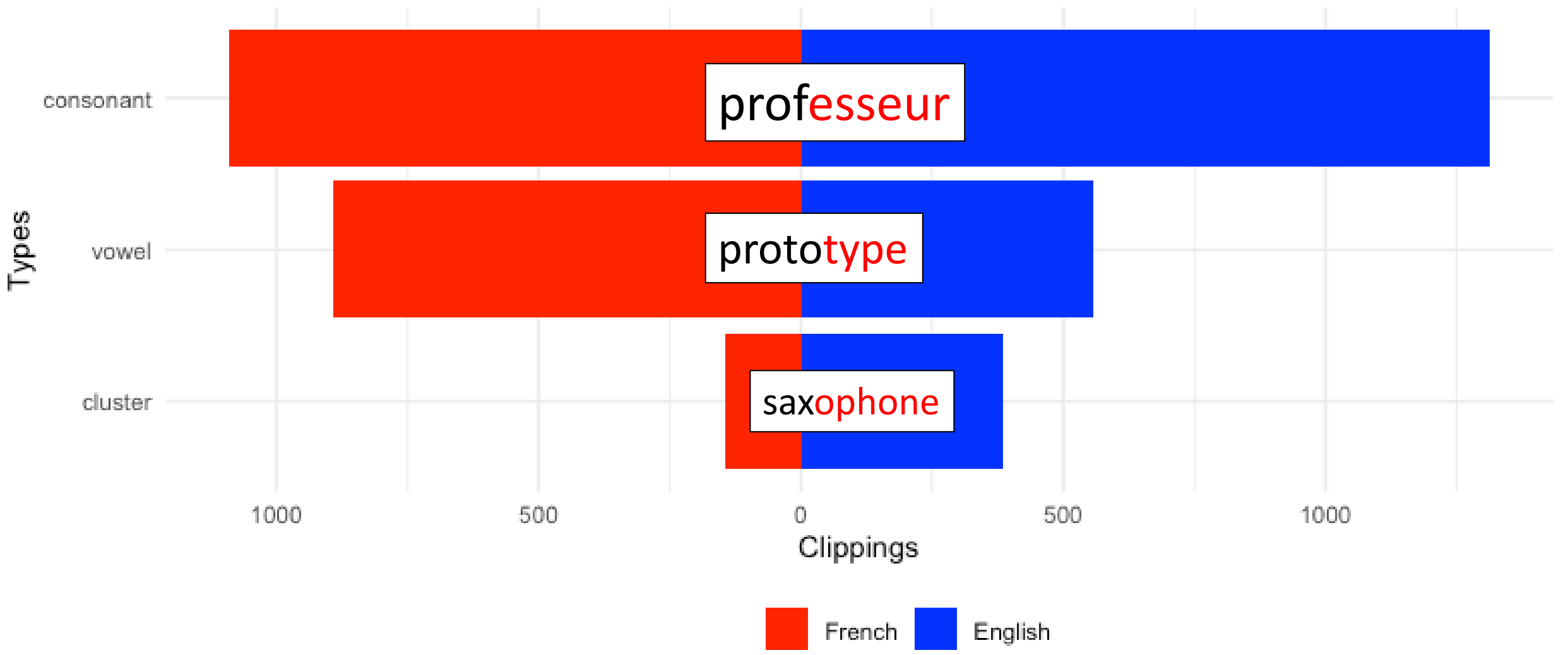
Clipping types



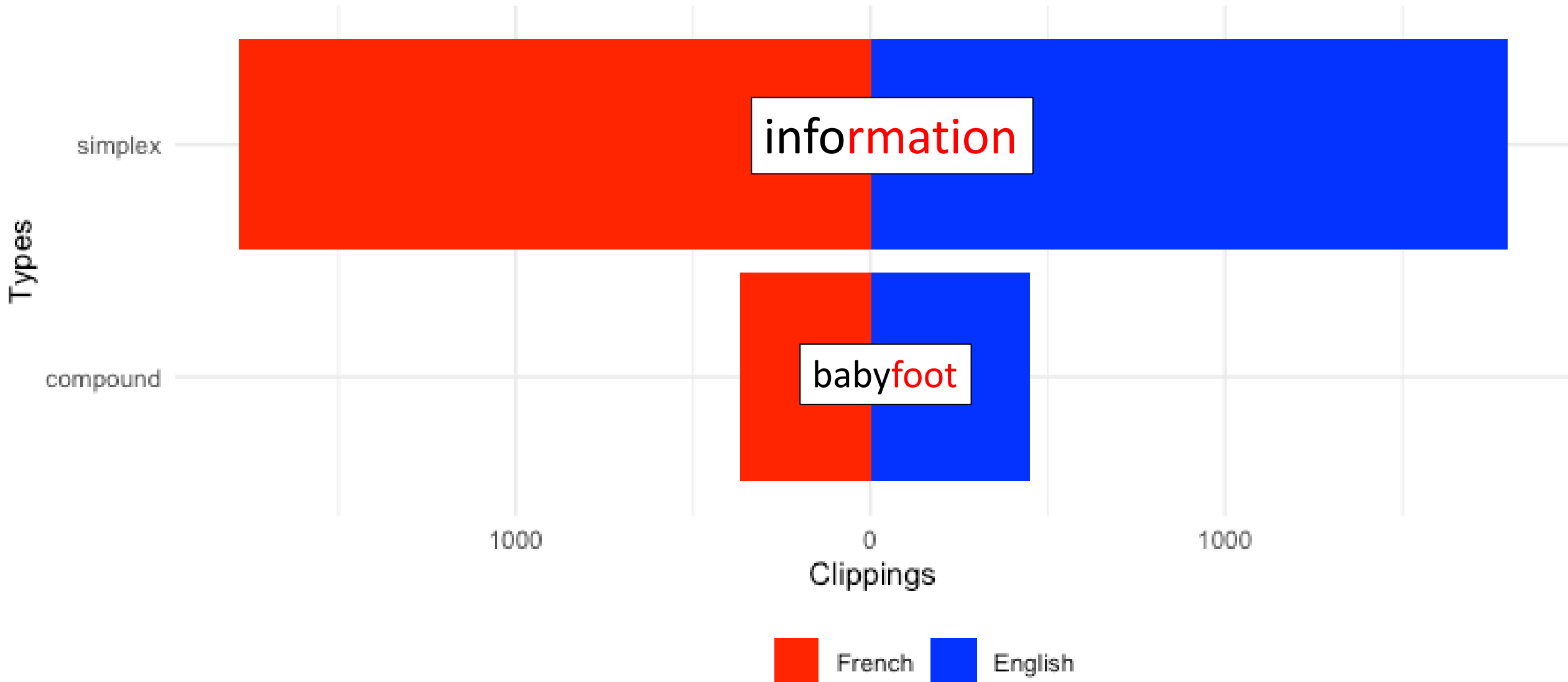
Morphological status



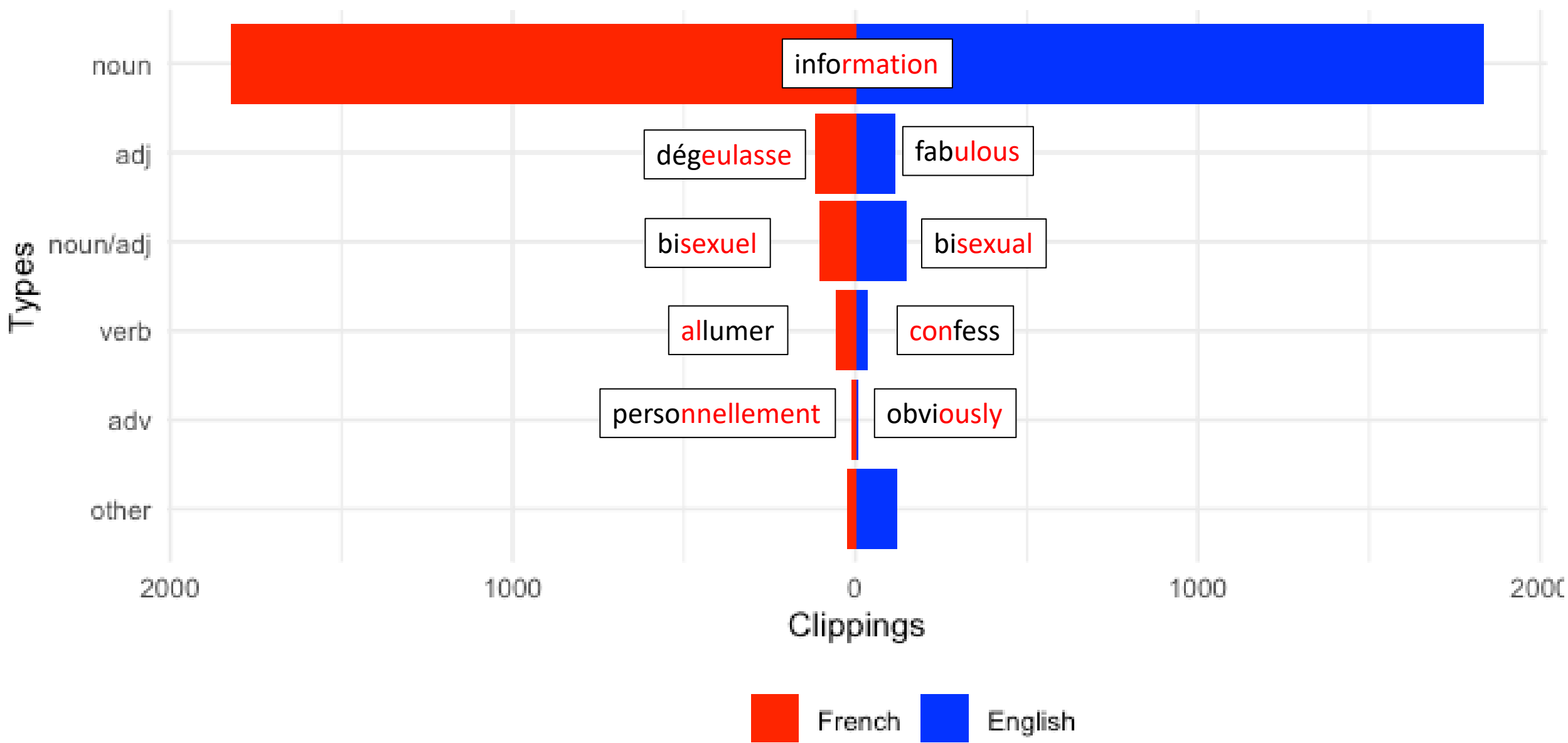
Final phonological segment



Compound source word



Part of speech



Variables in the clipping database

1. Number of syllables: 1, 2, 3, 4, ...
2. Clipped part: end, front, middle, middle-and-end, other
3. Morphological status: submorphemic, morphemic, word
4. Final sound: consonant, vowel, consonant cluster
5. Source word is a compound: yes, no
6. Part of speech: noun, adjective, noun/adjective, verb, adverb, other

More variables (not for this talk)

- Sound difference between clipping and source word: yes, no
- Semantic domain of the clipping: science, animals, drug use, ...
- Normalized frequency of the form of the clipping
- Orthographic neighborhood density of the clipping
- Dispersion of the clipping in corpus data (deviation of proportions)
- Number of word types that contain the clipping as a substring

Cross-tabulation of variables

- Do end-clipping and front-clipping yield clippings that differ in their morphological status?
- Is end-clipping more likely to produce clippings that correspond to a morpheme?

English clippings: Observed (and expected) frequencies

	Submorphemic	Morphemic	Word
End-clipping	1186 (1207.48)	325 (280.31)	234 (257.21)
Front-clipping	120 (155.00)	26 (35.98)	78 (33.02)
Middle-clipping	133 (110.02)	9 (25.54)	17 (23.44)
Middle-and-end-clipping	91 (65.05)	2 (15.10)	1 (13.86)
Other	38 (30.45)	2 (7.07)	4 (6.49)

Overrepresented

paleodiet

	Submorphemic	Morphemic	Word
End-clipping	1186 (1207.48)	325 (280.31)	234 (257.21)
Front-clipping	120 (155.00)	26 (35.98)	78 (33.02)
Middle-clipping	133 (110.02)	9 (25.54)	17 (23.44)
Middle-and-end-clipping	91 (65.05)	2 (15.10)	1 (13.86)
Other	38 (30.45)	2 (7.07)	4 (6.49)

romantic
comedy

nightclub

Cross-tabulation of variables

- Do end-clipping and front-clipping yield clippings that differ in their final phonological segment?
- Is end-clipping more likely to produce clippings that end in a vowel?

English clippings: Observed and expected frequencies

	Cluster	Consonant	Vowel
End-clipping	247 (296.98)	1022 (1008.35)	463 (426.67)
Front-clipping	30 (37.89)	154 (128.66)	37 (54.44)
Middle-clipping	80 (27.26)	62 (92.57)	17 (39.17)
Middle-and-end-clipping	17 (17.32)	60 (58.80)	24 (24.88)
Other	13 (7.54)	16 (25.62)	15 (10.84)

Overrepresented

quake, shroom

dino, limo

	Cluster	Consonant	Vowel
End-clipping	247 (296.98)	1022 (1008.35)	463 (426.67)
Front-clipping	30 (37.89)	154 (128.66)	37 (54.44)
Middle-clipping	80 (27.26)	62 (92.57)	17 (39.17)
Middle-and-end-clipping	17 (17.32)	60 (58.80)	24 (24.88)
Other	13 (7.54)	16 (25.62)	15 (10.84)

shorts, stats, gents, nugs

Configural Frequency Analysis

1. Number of syllables: 1, 2, 3, 4
2. Clipped part: end, front, middle, middle-and-end, other
3. Morphological status: submorphemic, morphemic, word
4. Final sound: consonant, consonant cluster, vowel
5. Source word is a compound: yes, no
6. Part of speech: noun, adjective, noun/adjective, verb, adverb, other

Configural Frequency Analysis

1. Number of syllables: **1**, 2, 3, 4
2. Clipped part: **end**, front, middle, middle-and-end, other
3. Morphological status: **submorphemic**, morphemic, word
4. Final sound: **consonant**, consonant cluster, vowel
5. Source word is a compound: yes, **no**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

Configural Frequency Analysis

1. Number of syllables: **1**, 2, 3, 4
2. Clipped part: **end**, front, middle, middle-and-end, other
3. Morphological status: **submorphemic**, morphemic, word
4. Final sound: **consonant**, consonant cluster, vowel
5. Source word is a compound: yes, **no**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

crocodile

English

Observed: 438

Expected: 206

Configural Frequency Analysis

1. Number of syllables: **1**, 2, 3, 4
2. Clipped part: **end**, front, middle, middle-and-end, other
3. Morphological status: **submorphemic**, morphemic, word
4. Final sound: **consonant**, consonant cluster, vowel
5. Source word is a compound: yes, **no**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

crocodile

2. What are the similarities and differences between English and French?

Configural Frequency Analysis

1. Number of syllables: 1, 2, 3, 4
2. Clipped part: end, front, middle, middle-and-end, other
3. Morphological status: submorphemic, morphemic, word
4. Final sound: consonant, consonant cluster, vowel
5. Source word is a compound: yes, no
6. Part of speech: noun, adjective, noun/adjective, verb, adverb, other
7. Language: English, French

Analysis

- Are there clipping configurations that are shared across both languages?
- Are there configurations that are language-specific?
- Are there asymmetries, so that a preferred configuration in English is a dispreferred configuration in French, or vice versa?

Shared preferences
and dispreferences

Shared type “croc”

English

Observed: 438

Expected: 206

French

Observed: 339

Expected: 198

1. Number of syllables: **1**, 2, 3, 4
2. Clipped part: **end**, front, middle, middle-and-end, other
3. Morphological status: **submorphemic**, morphemic, word
4. Final sound: **consonant**, consonant cluster, vowel
5. Source word is a compound: yes, **no**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

crocodile, **professeur**, **démission**, ...

Shared type “klepto”

English
Observed: 127
Expected: 34

French
Observed: 165
Expected: 32

1. Number of syllables: 1, **2**, 3, 4
2. Clipped part: **end**, front, middle, middle-and-end, other
3. Morphological status: submorphemic, **morphemic**, word
4. Final sound: consonant, consonant cluster, **vowel**
5. Source word is a compound: yes, **no**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

klepto**maniac**, chemo**therapy**, mult**inational**, ...

Shared type “romcom”

English
Observed: 32
Expected: 2

French
Observed: 21
Expected: 2

1. Number of syllables: 1, **2**, 3, 4
2. Clipped part: end, front, middle, **middle-and-end**, other
3. Morphological status: **submorphemic**, morphemic, word
4. Final sound: **consonant**, consonant cluster, vowel
5. Source word is a compound: no, **yes**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

rom**antic** comed**y**, devo**ir** suppl**émentaire**, ...

Shared antitype “psy”

English
Observed: 31
Expected: 125

French
Observed: 26
Expected: 120

1. Number of syllables: **1**, 2, 3, 4
2. Clipped part: **end**, front, middle, middle-and-end, other
3. Morphological status: **submorphemic**, morphemic, word
4. Final sound: consonant, consonant cluster, **vowel**
5. Source word is a compound: yes, **no**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

psychiatre, cucumber, detective, ...

French preferences and dispreferences

French type “hebdo”

French
Observed: 287
Expected: 114

English
Observed: 96
Expected: 119
n.s.

1. Number of syllables: 1, **2**, 3, 4
2. Clipped part: **end**, front, middle, middle-and-end, other
3. Morphological status: **submorphemic**, morphemic, word
4. Final sound: consonant, consonant cluster, **vowel**
5. Source word is a compound: yes, **no**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

hebdom**adaire**, ciné**ma**, restaur**ant**, ...

French type “toxico”

French
Observed: 55
Expected: 6

English
Observed: 15
Expected: 6
n.s.

1. Number of syllables: 1, 2, **3**, 4
2. Clipped part: **end**, front, middle, middle-and-end, other
3. Morphological status: submorphemic, **morphemic**, word
4. Final sound: consonant, consonant cluster, **vowel**
5. Source word is a compound: yes, **no**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

toxic**mane**, bibli**thèque**, magnéto**scope**, ...

French antitype “deuche”

French
Observed: 11
Expected: 45

English
Observed: 31
Expected: 46
n.s.

1. Number of syllables: **1**, 2, 3, 4
2. Clipped part: **end**, front, middle, middle-and-end, other
3. Morphological status: **submorphemic**, morphemic, word
4. Final sound: **consonant**, consonant cluster, vowel
5. Source word is a compound: **yes**, no
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other



deux-che**vaux**, sous-mar**in**, défaut de pai**ement**, ...
public **house**, coke-a-co**la**, federal **agent**, ...

English preferences and dispreferences



GLOSSA
a journal of general linguistics

Hilpert et al. 2023

English type “specs”

English
Observed: 37
Expected: 2

French
Observed: 2
Expected: 2
n.s.

1. Number of syllables: **1**, 2, 3, 4
2. Clipped part: end, front, **middle**, middle-and-end, other
3. Morphological status: **submorphemic**, morphemic, word
4. Final sound: consonant, **consonant cluster**, vowel
5. Source word is a compound: yes, **no**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

spect**acles**, circum**stances**, nug**gets**, ...

English anti-type “rehab”

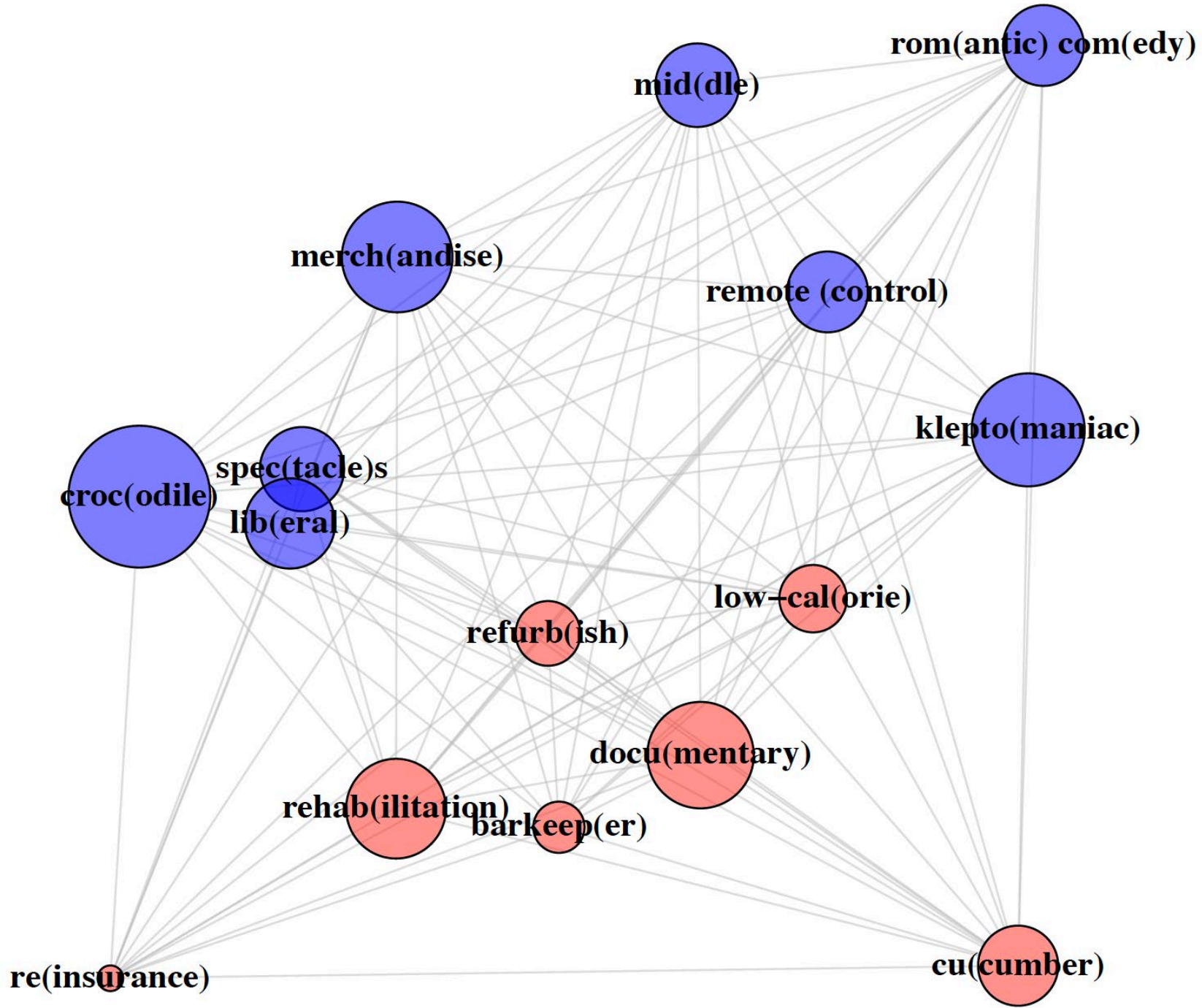
English
Observed: 72
Expected: 158

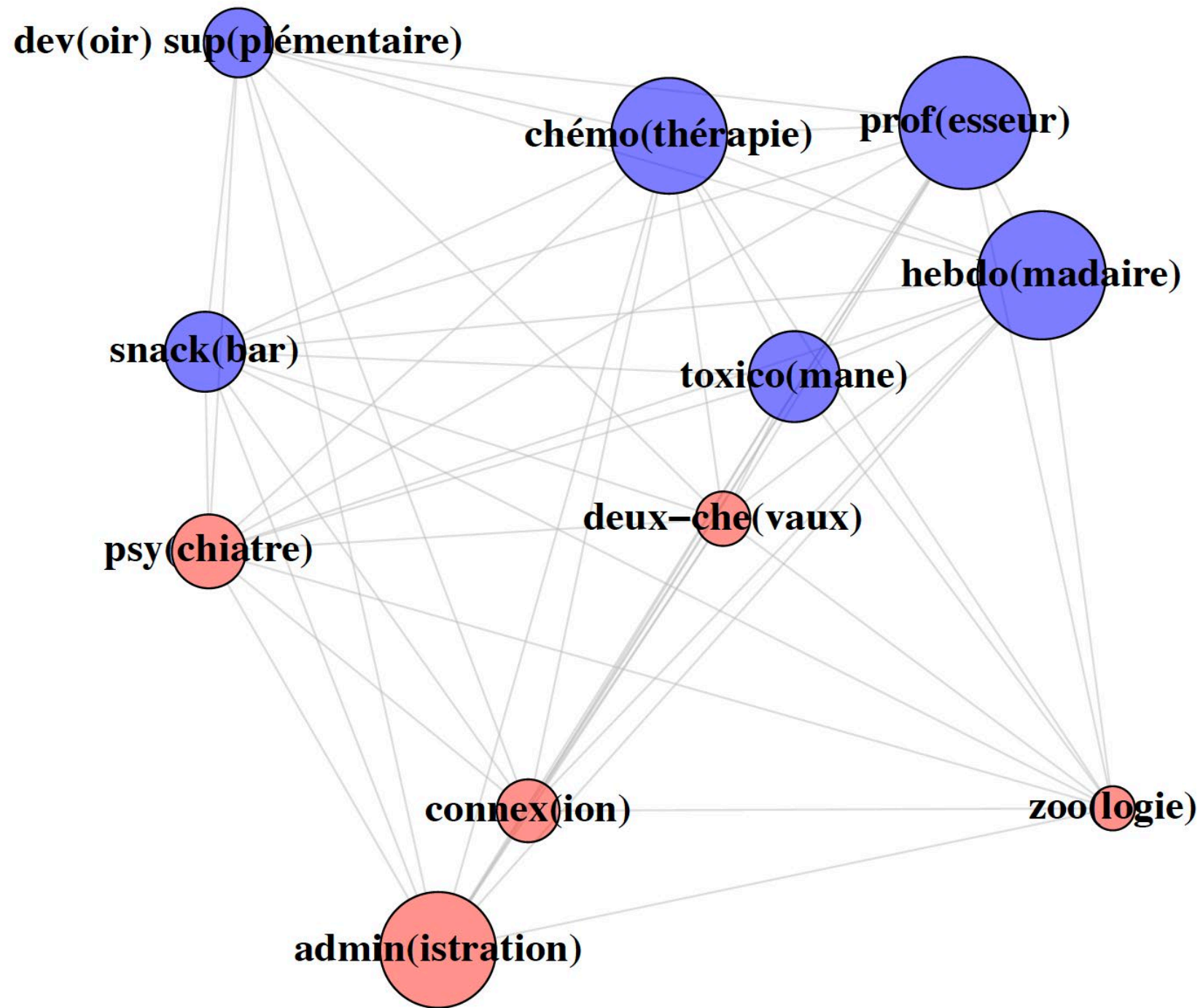
French
Observed: 165
Expected: 189
n.s.

1. Number of syllables: 1, **2**, 3, 4
2. Clipped part: **end**, front, middle, middle-and-end, other
3. Morphological status: **submorphemic**, morphemic, word
4. Final sound: **consonant**, consonant cluster, vowel
5. Source word is a compound: yes, **no**
6. Part of speech: **noun**, adjective, noun/adjective, verb, adverb, other

rehab**ilitation**, intellig**ence**, administrat**ion**, ...

Networks of constructions





4. Discussion and conclusions

Clippings in English and French

- In both languages, clipping is a highly productive word formation process that exhibits variation on multiple levels:
 - length, segmental phonology, morphology, parts of speech, ...
- There are strong parallels:
 - noun bias, preference for end-clipping, submorphemic clippings
- We observe a number of differences:
 - French prefers disyllabic clippings, English prefers monosyllabic clippings.
 - More morphemic clippings in French, more whole-word clippings in English
 - Middle-clippings and middle-and-end clippings are relatively more common in English.

Shared preferences and dispreferences

- Similarities

- Configurations such as the 'croc' type, the 'klepto' type, and the 'dev sup' type are preferred in both English and French.
- Monosyllabic vowel-final clippings are dispreferred in both languages.
- Disyllabic consonant-final clippings are dispreferred in both languages.

- Differences

- French types 'hebdo' and 'toxico' are not as strongly preferred in English.
- The French anti-type 'deuche' is not dispreferred in English.
- The English type 'specs' is not preferred in French.
- The English anti-type 'rehab' is less strongly dispreferred in French.

General observations

- Clippings are optimized for recoverability.
- Clippings of different lengths differ in their phonological and morphological characteristics.
- There is not one single clipping schema, but there are different statistically significant types that speakers orient to when they form new clippings.
- Some of these types are identical or highly similar across English and French, others are language-specific.



David Correia Saavedra

Many thanks!



Jennifer Rains

martin.hilpert@unine.ch

david.correiasaavedra@unine.ch

jennifer.rains@unine.ch



deuche