

Syllable structure and phonotactics

LING 20: Introduction to Linguistic Analysis

UCLA · Winter 2022

Syllables

- When we pronounce words like ‘understanding’, we do not just pronounce them as a simple string of sounds.
- Rather, the sounds are grouped into **SYLLABLES**.
- Syllable boundaries are indicated in the IPA with a period [.]:

Syllables

- When we pronounce words like ‘understanding’, we do not just pronounce them as a simple string of sounds.
- Rather, the sounds are grouped into **SYLLABLES**.
- Syllable boundaries are indicated in the IPA with a period [.]:
 - [ʌndɹ̩stændɪŋ]
 - [ɹɛkɪə̯ʃən]
 - [bʌlɔ̯wni]

Syllables

- When we pronounce words like ‘understanding’, we do not just pronounce them as a simple string of sounds.
- Rather, the sounds are grouped into **SYLLABLES**.
- Syllable boundaries are indicated in the IPA with a period [.]:
 - [ʌn.dɪ̯.stæn.dɪŋ]
 - [ɹɛ.kɪ.ə̃.jən]
 - [bʌ.ləʊ.ni]

Syllables are fixed

Native speakers of a language have clear intuitions about where syllable breaks **can** and **cannot** occur:

- (1) a. [ʌn.dɪ.ʌ.stæn.dɪŋ]
b. *[ʌ.ndɪ.ʌ.tænd.ɪŋ]

Syllables are fixed

Native speakers of a language have clear intuitions about where syllable breaks **can** and **cannot** occur:

(1) a. [ʌn.dɪ.stæn.dɪŋ]

b. *[ʌ.ndɪs.tænd.ɪŋ]

(2) a. [ɹɛ.kɪ.ə̃.jən]

b. *[ɹɛk.ɪ.ə̃.jən]

Syllables are fixed

Native speakers of a language have clear intuitions about where syllable breaks **can** and **cannot** occur:

(1) a. [ʌn.dɪ.stæn.dɪŋ]

b. *[ʌ.ndɪs.tænd.ɪŋ]

(2) a. [ɹɛ.kɪ.ə̃.jən]

b. *[ɹɛk.ɪ.ə̃]jən]

(3) a. [bʌ.ləʊ̃.ni]

b. *[bʌl.əʊ̃n.i]

Syllables are rule-governed

Question:

Is it possible that speakers have simply memorized the syllable breaks for all the words of a language?

Syllables are rule-governed

Question:

Is it possible that speakers have simply memorized the syllable breaks for all the words of a language?

Answer:

No! Speakers have clear intuitions about syllable breaks in words that they have never seen before and even nonce words (strings of sounds that could be words in a language but happen not to be).

Syllables are rule-governed

[bōwlgæstænd]

Syllables are rule-governed

[bōwlgæstænd]



[bōwl.ge.stænd]

Syllables are rule-governed

Conclusion:

The placement of syllables is not memorized, but instead is the result of a regular process, i.e. a **RULE**.

Syllables are rule-governed

Conclusion:

The placement of syllables is not memorized, but instead is the result of a regular process, i.e. a **RULE**.

Refresher:

When you know a language, you know:

1. a finite set of **basic elements**
2. a set of **rules** for combining basic elements (i.e. a **grammar**)

Syllables are rule-governed

Conclusion:

The placement of syllables is not memorized, but instead is the result of a regular process, i.e. a **RULE**.

Refresher:

When you know a language, you know:

1. a finite set of **basic elements**
→ *What sounds does a word consist of?*
2. a set of **rules** for combining basic elements (i.e. a **grammar**)
→ *How are sounds arranged into syllables?*

What is the **mechanism** that groups sounds
into syllables?

What is the **mechanism** that groups sounds into syllables?

To answer this question, we will need to learn a bit about the **structure of syllables** and the **phonotactics** of English.

The structure of syllables

Every syllable is made up of (at most) three basic parts:

Terminology: Onset

The **ONSET** is the consonant(s) at the front of the syllable.

Terminology: Nucleus

The **NUCLEUS** is the sound (typically, a vowel) in the middle of the syllable.

Terminology: Coda

The **CODA** is the consonant(s) at the end of the syllable.

The structure of syllables

Not all of these parts are present in all syllables:

- Some syllables have no onset, e.g. [it].
- Some syllables have no coda, e.g. [ti].
- However, every syllable does have a nucleus.

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|-------|-------|---------|------|
| [bæn] | | | |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|--------|--------------|----------------|-------------|
| [bæ̃n] | [b] | [æ̃] | [n] |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|-------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | | | |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|-------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|-------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |
| [ɛg] | | | |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|-------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |
| [ɛg] | — | [ɛ] | [g] |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|---------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |
| [εg] | — | [ε] | [g] |
| [stænd] | | | |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|---------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |
| [ɛg] | — | [ɛ] | [g] |
| [stænd] | [st] | [æ] | [nd] |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|---------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |
| [ɛg] | — | [ɛ] | [g] |
| [stænd] | [st] | [æ] | [nd] |
| [bɔ̃] | | | |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|---------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |
| [ɛg] | — | [ɛ] | [g] |
| [stænd] | [st] | [æ] | [nd] |
| [bɔ̃] | [b] | [ɔ̃] | — |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|---------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |
| [ɛg] | — | [ɛ] | [g] |
| [stænd] | [st] | [æ] | [nd] |
| [bɔ̃] | [b] | [ɔ̃] | — |
| [spɹɛ̃] | | | |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|---------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |
| [ɛg] | — | [ɛ] | [g] |
| [stænd] | [st] | [æ] | [nd] |
| [bɔ̃] | [b] | [ɔ̃] | — |
| [spɾɐ̃] | [spɾ] | [ɐ̃] | — |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|---------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |
| [ɛg] | — | [ɛ] | [g] |
| [stænd] | [st] | [æ] | [nd] |
| [bɔ̃] | [b] | [ɔ̃] | — |
| [spɹɛ̃] | [spɹ] | [ɛ̃] | — |
| [ã] | | | |

Examples: Syllable structure

| | Onset | Nucleus | Coda |
|----------|-------|---------|------|
| [bæn] | [b] | [æ] | [n] |
| [bi] | [b] | [i] | — |
| [ɛg] | — | [ɛ] | [g] |
| [stænd] | [st] | [æ] | [nd] |
| [bɔ̃] | [b] | [ɔ̃] | — |
| [spɹɛ̃j] | [spɹ] | [ɛ̃] | — |
| [â] | — | [â] | — |

Rhymes

The nucleus and coda group together to the exclusion of the onset because syllables that have the same nucleus and coda stand in a special relationship to each other: they **RHyme**.

- [spɪt], [sɪt], and [kɪt] all rhyme.
- [sɪt], [sɪts], [sɪn], and [sɪp] do not rhyme.

The structure of syllables

Every syllable is made up of (at most) three basic parts:

Terminology: Onset

The **ONSET** is the consonant(s) at the front of the syllable.

Terminology: Nucleus

The **NUCLEUS** is the sound (typically, a vowel) in the middle of the syllable.

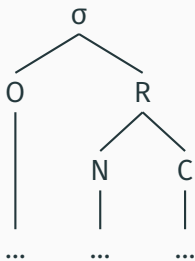
Terminology: Coda

The **CODA** is the consonant(s) at the end of the syllable.

Terminology: Rhyme

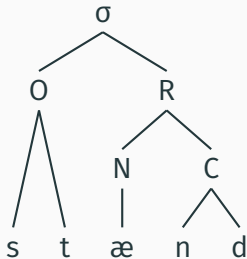
The **NUCLEUS** and **CODA**, if there is one.

Syllabification trees

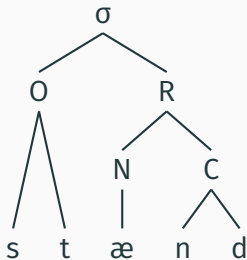


- **σ**: syllable
- **R**: rhyme
- **O**: onset
- **C**: coda
- **N**: nucleus

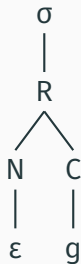
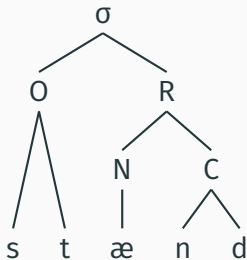
Examples: Syllabification trees



Examples: Syllabification trees



Examples: Syllabification trees



Building syllables

How do we decide which sounds become onsets, nuclei, and codas?

To answer this question, we need to learn a bit about **phonotactic constraints**.

Only one of the following strings could be an English word:

- (4) a. [pfluk]
b. [tloʔ.tli]
c. [tsa]
d. [tæg.nɪsp]
e. [nuktʃ]

Only one of the following strings could be an English word:

- (4) a. [pfluk]
b. [tloʔ.tli]
c. [tsa]
d. **[tæg.nɪsp]**
e. [nuktʃ]

Phonotactics

Terminology: Phonotactics

The rules that determine possible onsets and codas are called **PHONOTACTIC CONSTRAINTS** (or **PHONOTACTICS**).

- [pfl] is not a possible onset in English.
- [tl] is not a possible onset in English.
- [ts] is not a possible onset in English.
- [ktʃ] is not a possible coda in English.

Phonotactics crosslinguistically

Different languages have different phonotactic constraints:

(5) **German (Germanic):**

[pfluk] 'plow'

(6) **Nahuatl (Uto-Aztecan; Mexico):**

[tloʔ.tli] 'hawk'

(7) **Tlingit (Na-Dene; Alaska):**

a. [tsa] 'seal'

b. [nuktʃ] 'to do'

English: The most complex onsets

| | | | | |
|-----|-----|-----|---------------|--|
| [s] | [p] | [l] | [splɪn] | 'spleen' |
| | | [ɹ] | [spɹeɪ] | 'spray' |
| | | [j] | [spju] | 'spew' |
| | [t] | [j] | [stju] | 'stew' (<i>some varieties</i>) |
| | | [ɹ] | [stɹɪŋ] | 'string' (<i>for some, [t] → [tʃ]</i>) |
| | [k] | [w] | [skwɒʃ] | 'squash' |
| | | [j] | [skju] | 'skew' |
| | | [l] | [skliːrɒwsɪs] | 'sclerosis' |
| | | [ɹ] | [skɹeɪp] | 'scrape' |

Phonotactics and syllabification

[ʌn.dɪ.stæn.dɪŋ] vs. *[ʌ.ndɪs.tænd.ɪŋ]

Phonotactics and syllabification

[ʌn.dɪ.stæn.dɪŋ] vs. *[ʌ.ndɪs.tænd.ɪŋ]

Phonotactic constraints guide how a word is syllabified:

→ [nd] is not a possible onset in English.

Phonotactics and syllabification

- **How to tell whether a sequence of sounds is a possible onset or coda?**

- A string is a legal **ONSET** if it can begin a word
- A string is a legal **CODA** if it can end a word

- Try:

| | | |
|-----|-----|----|
| pl | bz | bd |
| slk | sts | lk |
| h | ŋ | θ |
| ɹb | pt | sk |

The syllabification algorithm

1. All vowels become **nuclei**. (NB: Diphthong → one nucleus)

The syllabification algorithm

1. All vowels become **nuclei**. (NB: Diphthong → one nucleus)
2. For each nucleus, find the **largest** continuous sequence immediately preceding the nucleus and allowed by the language's phonotactics to be an **onset**; make it an onset.

The syllabification algorithm

1. All vowels become **nuclei**. (NB: Diphthong → one nucleus)
2. For each nucleus, find the **largest** continuous sequence immediately preceding the nucleus and allowed by the language's phonotactics to be an **onset**; make it an onset.
3. For each nucleus, find the largest continuous sequence (i) immediately following the nucleus, (ii) not part of an onset, and (iii) allowed by the language's phonotactics to be a **coda**; make it a coda.

The syllabification algorithm

1. All vowels become **nuclei**. (NB: Diphthong → one nucleus)
2. For each nucleus, find the **largest** continuous sequence immediately preceding the nucleus and allowed by the language's phonotactics to be an **onset**; make it an onset.
3. For each nucleus, find the largest continuous sequence (i) immediately following the nucleus, (ii) not part of an onset, and (iii) allowed by the language's phonotactics to be a **coda**; make it a coda.
4. For each nucleus, make a **rhyme** out of the nucleus and, if there is one, the immediately following coda.

The syllabification algorithm

1. All vowels become **nuclei**. (NB: Diphthong → one nucleus)
2. For each nucleus, find the **largest** continuous sequence immediately preceding the nucleus and allowed by the language's phonotactics to be an **onset**; make it an onset.
3. For each nucleus, find the largest continuous sequence (i) immediately following the nucleus, (ii) not part of an onset, and (iii) allowed by the language's phonotactics to be a **coda**; make it a coda.
4. For each nucleus, make a **rhyme** out of the nucleus and, if there is one, the immediately following coda.
5. For each rhyme, make a **syllable** out of the rhyme and, if there is one, the immediately preceding onset.

Syllabification: Beyond phonotactics

Observation:

There is more to syllabification than just phonotactics:

[bʌ.ləʊ̌n.i] vs. *[bʌl.əʊ̌n.i]

Syllabification: Beyond phonotactics

Observation:

There is more to syllabification than just phonotactics:

[bʌ.ləʊ̯.ni] vs. *[bʌ.l.əʊ̯n.i]

Question:

Both syllabifications respect the phonotactic constraints of English, yet only the one on the left is correct. Why?

Onsets before codas

The answer to this question comes from the syllabification algorithm itself:

- The algorithm determines onsets before codas.
- Therefore, if a sound or combination of sounds could in principle be either an onset or a coda, it will be an onset.

→ **Only sounds that cannot become onsets become codas.**

Aside: Syllabic consonants

- In some cases, the sounds [l], [n], [m], and [ɹ] can be nuclei in English:
 - (8) a. kɪt̩ 'kitten'
 - b. kʌd̩ 'cuddle'
 - c. ɹɪð̩ 'rhythm'
 - d. bæɳ̩ 'banner'
- A more complicated syllabification algorithm is necessary for these cases, which we will not develop in this class.

Constraints on syllable shapes

English allows a fair number of possible syllable shapes because it allows onsets and codas to consist of multiple consonants.

Possible syllable shapes in English

V [ʌ]

VC [æt]

VCC [æsk]

VCCC [æskt]

Possible syllable shapes in English

| | | | |
|------|--------|-------|---------|
| V | [ʌ] | CV | [nôw] |
| VC | [æt] | CVC | [nat] |
| VCC | [æsk] | CVCC | [ɹæmp] |
| VCCC | [æskt] | CVCCC | [ɹæmps] |

Possible syllable shapes in English

| | | | | | |
|------|--------|-------|---------|--------|----------|
| V | [ʌ] | CV | [nōw] | CCV | [flu] |
| VC | [æt] | CVC | [nat] | CCVC | [flut] |
| VCC | [æsk] | CVCC | [ɹæmp] | CCVCC | [fluts] |
| VCCC | [æskt] | CVCCC | [ɹæmps] | CCVCCC | [kɹæfts] |

Possible syllable shapes in English

| | | | | | | | |
|------|--------|-------|---------|--------|----------|---------|-----------|
| V | [ʌ] | CV | [nōw] | CCV | [flu] | CCCV | [spɹi] |
| VC | [æt] | CVC | [nat] | CCVC | [flut] | CCVC | [splin] |
| VCC | [æsk] | CVCC | [ɹæmp] | CCVCC | [fluts] | CCCVCC | [stɹɛŋθ] |
| VCCC | [æskt] | CVCCC | [ɹæmps] | CCVCCC | [kɹæfts] | CCCVCCC | [stɹɛŋθs] |

Possible syllable shapes crosslinguistically

Some languages have stronger constraints on possible syllable shapes:

- **Hawaiian (Polynesian):** CV, V
- **Indonesian (Malayic):** CV, V, VC, CVC
- **Hebrew (Semitic):** CV, CCV, CCVC, CVC, CVCC

Possible syllable shapes crosslinguistically

Other languages are more permissive than English in the syllable shapes that they allow:

(9) **Russian (Slavic):**

- a. [vzdrognʊtʲ] 'to flinch'
- b. [fskrʲɪtʲ] 'to unseal'
- c. [fsxrapʲɪvətʲ] 'to snort'

Possible syllable shapes crosslinguistically

Other languages are more permissive than English in the syllable shapes that they allow:

(9) **Russian (Slavic):**

- a. [vzdrognʊtʲ] 'to flinch'
- b. [fskrʲɪtʲ] 'to unseal'
- c. [fsxrapʲɪvətʲ] 'to snort'

(10) **Georgian (Kartvelian):**

- a. [brdyvna] 'to fight'
- b. [gvprckvnɪs] 'he is peeling us'

Differences in syllabification

Languages differ in how they group sounds into syllables:

(11) a. **English:**

[mɪ.stə.ɹi]

b. **Indonesian:**

[mis.te.ri]

This is because Indonesian phonotactics prohibit onsets that consist of more than one sound, while English allows them.

Differences in syllabification

- Every language will make onsets as large as possible.
- If a sound can be either an onset or a coda, every language will make it an onset.

Differences in syllabification

- Every language will make onsets as large as possible.
- If a sound can be either an onset or a coda, every language will make it an onset.

- **Consequence:**

The syllabification algorithm seems to be **universal**.

Differences in syllable structure result from differences in phonotactic constraints.

Differences in syllabification

Farsi: What are the possible syllable shapes?

| | | | |
|-------|---------------|--------|------------|
| [u] | 'she/he/they' | [ma] | 'we' |
| [an] | 'it' | [xub] | 'good' |
| [æsr] | 'afternoon' | [lotf] | 'kindness' |

Differences in syllabification

Farsi: What are the possible syllable shapes?

| | | | |
|-------|---------------|--------|------------|
| [u] | 'she/he/they' | [ma] | 'we' |
| [an] | 'it' | [xub] | 'good' |
| [æsr] | 'afternoon' | [lotf] | 'kindness' |

How would you syllabify these words in Farsi?

| | | | |
|-------|--------|---------|--------|
| xaste | nistam | gorosne | mamnum |
|-------|--------|---------|--------|

Phonotactics and borrowing

If a language borrows a word from another language that violates a phonotactic constraint of that language, this word will be modified to avoid this violation.

- '*Ptolemy*': [taləmi]
from Ancient Greek [ptolemaios]

Phonotactics and borrowing

If a language borrows a word from another language that violates a phonotactic constraint of that language, this word will be modified to avoid this violation.

- '*Ptolemy*': [taləmi]
from Ancient Greek [ptolemaios]
- '*gnostic*': [nastɪk]
from Ancient Greek [gnōstikos]

Phonotactics and borrowing

If a language borrows a word from another language that violates a phonotactic constraint of that language, this word will be modified to avoid this violation.

- '*Ptolemy*': [taləmi]
from Ancient Greek [ptolemaios]
- '*gnostic*': [nastɪk]
from Ancient Greek [gnōstikos]
- '*psychology*': [sāʝkalədʒi]
from Ancient Greek [psykēlogia]

Phonotactics and borrowing

If a language borrows a word from another language that violates a phonotactic constraint of that language, this word will be modified to avoid this violation.

- '*Ptolemy*': [taləmi]
from Ancient Greek [ptolemaios]
- '*gnostic*': [nastɪk]
from Ancient Greek [gnōstikos]
- '*psychology*': [sāʝkalədʒi]
from Ancient Greek [psykēlogia]
- '*tsunami*': [sunami]
from Japanese [tsuunami]

English → Japanese

- *'Birth control'* → [basu kontororuw]
- *'McDonalds'* → [makuɔdonalɯdo]

Summary:

- **Sounds are grouped into syllables, which have internal structure: onset, nucleus, coda, and rhyme.**
- **There is a deterministic algorithm that assigns syllable structure to strings of sounds.**
- **Phonotactic constraints play a key role.**
- **The syllabification algorithm seems to be universal.**
- **Different languages allow different syllable shapes because they differ in their phonotactics.**