The evolution of linguistic regularities and exceptions
Kenny Smith
kenny.smith@ed.ac.uk
Communication is widespread, but language is unique.
Language’s open-ended communicative power comes from its regular compositional structure
John loves Mary
Yesterday I saw a pink cat!
A Russian woman was sentenced to 21 years for trying to kill her doppelganger with a poisoned cheesecake!!!
Language is transmitted via repeated **learning** and **use**

Language is shaped by these processes

These processes create the regular compositional structure that makes language so powerful

And we can understand **exceptions** to those regularities (within and across languages) in terms of those same processes

From Smith, 2022, *Current Directions in Psychological Science* (based on e.g. Hurford, 1990, in *Logical Issues in Language Acquisition*)
Computational experiments
See e.g. Kirby, Griffiths, & Smith (2014). Current Opinion in Neurobiology for review
Lab experiments
See e.g. Smith (2022). *Current Directions in Psychological Science* for review
Artificial Language Learning
Communicative interaction
Iterated learning
Talk outline

Part 1: learning and use create regular compositional structure

Part 2: learning also explains where irregularity appears within languages (high-frequency things are more likely to be irregular)

Part 3: learning also explains where irregularity appears across languages (languages with fewer non-native speakers are more complex and irregular)
Part 1: learning and use create regular compositional structure

Simon Kirby (Edinburgh)

Monica Tamariz (Heriot Watt)

Hannah Cornish (Edinburgh)
Learning only

Kirby, Cornish & Smith, 2008, PNAS
(also e.g. Silvey, Kirby & Smith, 2015, Cog Sci)
Starting point: non-compositional ‘language’

wimaku → miniki → gepinini

kinimapi → wikuki → kikumi

pikuhemi → kimaki → pimikihe
Result: loss of distinctions, degeneracy

Kirby, Cornish & Smith, 2008, PNAS
Result: loss of distinctions, degeneracy

Kirby, Cornish & Smith, 2008, PNAS
Result: preservation of non-compositional system

Kirby, Tamariz, Cornish & Smith, 2015, Cognition
Kirby, Tamariz, Cornish & Smith, 2015, Cognition (also Winters, Kirby & Smith, 2015, Language & Cognition; Saldana, Kirby, Truswell & Smith , 2019, JoLE)
Result: (simple) compositional structure

Kirby, Tamariz, Cornish & Smith, 2015, *Cognition*
Summary of Part 1

Learning plus use leads to regular compositional structure
• Learning favours simplicity
• Communicative use prevents collapse to degeneracy
• To be both learnable and useful, be regular and compositional
Part 2: same processes explain where irregularity appears within languages

Smith, Ashton, & Sims-Williams, 2023, *Proc Cog Sci*
https://escholarship.org/uc/item/1mz1q97f
A Russian woman was sentenced to 21 years for trying to kill her doppelganger with a poisoned cheesecake!!!
But not all of language is regular and compositional

time, year, people, way, man, day, thing, child, work, life, ...
But not all of language is regular and compositional

time, year, people, way, man, day, thing, child, work, life, ...

is-was, has-had, does-did, says-said, gets-got, makes-made, goes-went, sees-saw, knows-knew, takes-took, thinks-thought, comes-came, gives-gave, looks-looked, ...
But not all of language is regular and compositional

time, year, people, way, man, day, thing, child, work, life, ...

is-was, has-had, does-did, says-said, gets-got, makes-made, goes-went, sees-saw, knows-knew, takes-took, thinks-thought, comes-came, gives-gave, looks-looked, ...

beat around the bush, let the cat out of the bag, bite the bullet, pull their leg, spill the beans, go cold turkey, ring a bell, give them a run for their money, find their feet
But not all of language is regular and compositional

time, year, people, way, man, day, thing, child, work, life, ...

is-was, has-had, does-did, says-said, gets-got, makes-made, goes-went, sees-saw, knows-knew, takes-took, thinks-thought, comes-came, gives-gave, looks-looked, ...

beat around the bush, let the cat out of the bag, bite the bullet, pull their leg, spill the beans, go cold turkey, ring a bell, give them a run for their money, find their feet, your bum’s out the window, ...
Proposed mechanisms for the frequency-irregularity correlation

Learning-based: frequent items are more resistant to pressures for regularity operating in learning (e.g. analogy)
• e.g. Bybee (1995); Sims-Williams (2022)

Usage-based: frequent items more susceptible to reductive sound change which causes irregularity
• e.g. Garrett (2015); Bybee (2017); Todd, Pierrehumbert, & Hay (2019)
Can we reproduce the frequency-irregularity correlation under controlled conditions in the lab?

Can we test mechanisms involved?
<table>
<thead>
<tr>
<th></th>
<th>Object 1</th>
<th>Object 2</th>
<th>Object 3</th>
<th>Object 4</th>
<th>Object 5</th>
<th>Object 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular image</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>Singular label</td>
<td>viza</td>
<td>drashru</td>
<td>wodra</td>
<td>mowo</td>
<td>shrunu</td>
<td>plonu</td>
</tr>
<tr>
<td>Plural image</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td>Plural label</td>
<td>zawo</td>
<td>drashrubli</td>
<td>huvi</td>
<td>mowosla</td>
<td>nuplo</td>
<td>plonuri</td>
</tr>
<tr>
<td>Object</td>
<td>Singular image</td>
<td>Singular label</td>
<td>Plural image</td>
<td>Plural label</td>
<td>Number of trials in uniform training</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------------</td>
<td>----------------</td>
<td>--------------</td>
<td>--------------</td>
<td>-------------------------------------</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><img src="image1.png" alt="Image 1" /></td>
<td>viza</td>
<td><img src="image2.png" alt="Image 2" /></td>
<td>drashru</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td><img src="image3.png" alt="Image 3" /></td>
<td>wodra</td>
<td><img src="image4.png" alt="Image 4" /></td>
<td>huvi</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td><img src="image5.png" alt="Image 5" /></td>
<td>mowo</td>
<td><img src="image6.png" alt="Image 6" /></td>
<td>mowosla</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td><img src="image7.png" alt="Image 7" /></td>
<td>shrunu</td>
<td><img src="image8.png" alt="Image 8" /></td>
<td>nuplo</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td><img src="image9.png" alt="Image 9" /></td>
<td>plonu</td>
<td><img src="image10.png" alt="Image 10" /></td>
<td>plonuri</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td><img src="image11.png" alt="Image 11" /></td>
<td></td>
<td><img src="image12.png" alt="Image 12" /></td>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Object 1</td>
<td>Object 2</td>
<td>Object 3</td>
<td>Object 4</td>
<td>Object 5</td>
<td>Object 6</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Singular image</td>
<td><img src="image1" alt="Image" /></td>
<td><img src="image2" alt="Image" /></td>
<td><img src="image3" alt="Image" /></td>
<td><img src="image4" alt="Image" /></td>
<td><img src="image5" alt="Image" /></td>
<td><img src="image6" alt="Image" /></td>
</tr>
<tr>
<td>Singular label</td>
<td>viza</td>
<td>drashru</td>
<td>wodra</td>
<td>mowo</td>
<td>shrunu</td>
<td>plonu</td>
</tr>
<tr>
<td>Plural image</td>
<td><img src="image1_plural" alt="Image" /></td>
<td><img src="image2_plural" alt="Image" /></td>
<td><img src="image3_plural" alt="Image" /></td>
<td><img src="image4_plural" alt="Image" /></td>
<td><img src="image5_plural" alt="Image" /></td>
<td><img src="image6_plural" alt="Image" /></td>
</tr>
<tr>
<td>Plural label</td>
<td>zawo</td>
<td>drashrubli</td>
<td>huvi</td>
<td>mowosla</td>
<td>nuplo</td>
<td>plonuri</td>
</tr>
<tr>
<td>Number of trials in uniform training</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Number of trials in skewed training</td>
<td>24</td>
<td>24</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>
Prediction for uniform condition

<table>
<thead>
<tr>
<th></th>
<th>Initial (Generation 0)</th>
<th>...</th>
<th>Generation ∞</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mid frequency, suppletive</strong></td>
<td>zawo</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mid frequency, suffixing</strong></td>
<td>drashrubli</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mid frequency, suppletive</strong></td>
<td>huvi</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mid frequency, suffixing</strong></td>
<td>mowosla</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mid frequency, suppletive</strong></td>
<td>nuplo</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mid frequency, suffixing</strong></td>
<td>plonuri</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Prediction for uniform condition

Emergence of a single regular rule

<table>
<thead>
<tr>
<th>Initial (Generation 0)</th>
<th>...</th>
<th>Generation ∞</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mid frequency, suppletive</strong></td>
<td>zawo</td>
<td><strong>vizabli</strong></td>
</tr>
<tr>
<td><strong>Mid frequency, suffixing</strong></td>
<td>drashrubli</td>
<td><strong>drashrubli</strong></td>
</tr>
<tr>
<td><strong>Mid frequency, suppletive</strong></td>
<td>huvi</td>
<td><strong>wodrabli</strong></td>
</tr>
<tr>
<td><strong>Mid frequency, suffixing</strong></td>
<td>mowosla</td>
<td><strong>mowobli</strong></td>
</tr>
<tr>
<td><strong>Mid frequency, suppletive</strong></td>
<td>nuplo</td>
<td><strong>shrunubli</strong></td>
</tr>
<tr>
<td><strong>Mid frequency, suffixing</strong></td>
<td>plonuri</td>
<td><strong>plonubli</strong></td>
</tr>
</tbody>
</table>
### Prediction for skewed condition

<table>
<thead>
<tr>
<th>Initial (Generation 0)</th>
<th>…</th>
<th>Generation ∞</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High frequency, suppletive</strong></td>
<td>zawo</td>
<td></td>
</tr>
<tr>
<td><strong>High frequency, suffixing</strong></td>
<td>drashrubli</td>
<td></td>
</tr>
<tr>
<td><strong>Low frequency, suppletive</strong></td>
<td>huvi</td>
<td></td>
</tr>
<tr>
<td><strong>Low frequency, suffixing</strong></td>
<td>mowosla</td>
<td></td>
</tr>
<tr>
<td><strong>Low frequency, suppletive</strong></td>
<td>nuplo</td>
<td></td>
</tr>
<tr>
<td><strong>Low frequency, suffixing</strong></td>
<td>plonuri</td>
<td></td>
</tr>
</tbody>
</table>
Prediction for skewed condition

A single regular rule??

<table>
<thead>
<tr>
<th></th>
<th>Initial (Generation 0)</th>
<th>...</th>
<th>Generation ∞</th>
</tr>
</thead>
<tbody>
<tr>
<td>High frequency,</td>
<td></td>
<td></td>
<td>vizabli</td>
</tr>
<tr>
<td>suppletive</td>
<td></td>
<td></td>
<td>drashrubli</td>
</tr>
<tr>
<td>High frequency,</td>
<td></td>
<td></td>
<td>wodrabli</td>
</tr>
<tr>
<td>suffixing</td>
<td></td>
<td></td>
<td>mowobli</td>
</tr>
<tr>
<td>Low frequency,</td>
<td></td>
<td></td>
<td>shrunubli</td>
</tr>
<tr>
<td>suppletive</td>
<td></td>
<td></td>
<td>plonubli</td>
</tr>
<tr>
<td>Low frequency,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>suffixing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low frequency,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>suppletive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low frequency,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>suffixing</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

High frequency, suppletive: zawo
High frequency, suffixing: drashrubli
Low frequency, suppletive: huvi
Low frequency, suffixing: mowosla
Low frequency, suppletive: nuplo
Low frequency, suffixing: plonuri
Prediction for skewed condition

High frequency items remain exceptional?

<table>
<thead>
<tr>
<th>Initial (Generation 0)</th>
<th>...</th>
<th>Generation $\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High frequency, suppletive</strong></td>
<td>zawo</td>
<td>zawo</td>
</tr>
<tr>
<td><strong>High frequency, suffixing</strong></td>
<td>drashrubli</td>
<td>drashrubli</td>
</tr>
<tr>
<td><strong>Low frequency, suppletive</strong></td>
<td>huvi</td>
<td>wodrasla</td>
</tr>
<tr>
<td><strong>Low frequency, suffixing</strong></td>
<td>mowosla</td>
<td>mowosla</td>
</tr>
<tr>
<td><strong>Low frequency, suppletive</strong></td>
<td>nuplo</td>
<td>shrunusla</td>
</tr>
<tr>
<td><strong>Low frequency, suffixing</strong></td>
<td>plonuri</td>
<td>plonusla</td>
</tr>
</tbody>
</table>
Prediction for skewed condition

High frequency suppletive remains exceptional, high frequency suffix becomes the regular?
Prediction for skewed condition

High frequency suppletive remains exceptional, high frequency suffix becomes the regular?
Prediction for skewed condition

A single regular rule??

<table>
<thead>
<tr>
<th>High frequency, suppletive</th>
<th>Initial (Generation 0)</th>
<th>...</th>
<th>Generation $\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>zawo</td>
<td>1</td>
<td>vizablili</td>
</tr>
<tr>
<td>High frequency, suffixing</td>
<td>drashrubli</td>
<td>1</td>
<td>drashrubli</td>
</tr>
<tr>
<td>Low frequency, suppletive</td>
<td>huvi</td>
<td>1</td>
<td>wodrabi</td>
</tr>
<tr>
<td>Low frequency, suffixing</td>
<td>mowosla</td>
<td>1</td>
<td>mowobi</td>
</tr>
<tr>
<td>Low frequency, suppletive</td>
<td>nuplo</td>
<td>1</td>
<td>shrunubi</td>
</tr>
<tr>
<td>Low frequency, suffixing</td>
<td>plonuri</td>
<td>1</td>
<td>plonubi</td>
</tr>
</tbody>
</table>
Uniform condition

- Mid
- Low
- High, Suffixing
- High, Suppletive

Skewed condition

Class Size

Generation

Uniform condition:
- Mid
- Low
- High, Suffixing
- High, Suppletive

Skewed condition:
- Mid
- Low
- High, Suffixing
- High, Suppletive
Uniform condition

Skewed condition

Mid

Low

High, Suffixing

High, Suppletive

Class Size

Generation

Uniform condition

Skewed condition

Mid

Low

High, Suffixing

High, Suppletive

Class Size

Generation

Uniform condition

Skewed condition

Mid

Low

High, Suffixing

High, Suppletive

Class Size

Generation

Uniform condition

Skewed condition

Mid

Low

High, Suffixing

High, Suppletive

Class Size

Generation

Uniform condition

Skewed condition

Mid

Low

High, Suffixing

High, Suppletive

Class Size

Generation

Uniform condition

Skewed condition

Mid

Low

High, Suffixing

High, Suppletive

Class Size

Generation
In the Skewed condition

High-frequency suppletives resist regularization

High-frequency suffixes become the regular, attracting lower-frequency items to their class.
Reminder: Proposed mechanisms for the frequency-irregularity correlation

Learning-based: frequent items are more resistant to pressures for regularity operating in learning (e.g. analogy)
- e.g. Bybee (1995); Sims-Williams (2022)

Usage-based: frequent items more susceptible to reductive sound change which causes irregularity
- e.g. Garrett (2015); Bybee (2017); Todd, Pierrehumbert, & Hay (2019)
e.g. viza-zawo

e.g. drashru-drashrubli
Truncation examples:

- **Suppletion**
  - e.g. viza-
  - zawo

- **Suffix**
  - e.g. drashru-
  - drashru
  - bli

Learning accuracy of plural form:

- Low
- Mid
- High
e.g. viza-zawo

e.g. drashru-drashrubli

- e.g. drashru-drashrubli, viza-vizabli, wodra-wodrabli, shrunu-shrunubli, ...
e.g. viza-zawo
e.g. drashru-drashrubli

Learning accuracy of plural form
Summary of Part 2

• Things regularize (as we’d expect based on Part 1)
• High frequency items behave differently
  • High-frequency suppletives **resist** regularization, yielding a frequency-irregularity correlation
  • High-frequency suffixes **become the regular**
• High frequency provides (some) protection from the learning penalty associated with being idiosyncratic, and provides a niche where irregularity can persist
• In relation to part 1: learning pressure depends on frequency and therefore differs in different parts of the grammar
Part 3: same processes also explain where irregularity appears across languages

Smith, 2024, Proc Cog Sci
https://osf.io/preprints/psyarxiv/2kzd4
Languages differ in their complexity

Maddieson, 1984, *Patterns of Sounds*

Winter & Bentz, 2013, *LDC*

Zaslavsky et al, 2019, *Cognitive Neuropsychology*

Bickel & Nichols, 2013, in *WALS*
The Linguistic Niche Hypothesis: Languages spoken in larger populations, with more non-native speakers, tend to be simpler*

Potential mechanism: uncertainty in interaction

e.g. Atkinson et al., 2017, JoLE; Raviv et al., 2019, Proc Roy Soc; Feher et al., 2019, JML; Frank & Smith, 2020, Proc Cog Sci; Loy et al., 2020, Discourse Processes; ...
Potential mechanism: learning from heterogeneous sources; learning from non-native speakers

Atkinson et al., 2018, *Cognitive Science*; Berdicevskis & Semenuks, 2022, *PLOS ONE*; this talk
Experiment 1: establishing an experimental proxy for native vs non-native speakers
Target language
**Training** (18 trials)

- *wona grolo wooshan*

**Testing** (18 trials)

- *sumuk snapop boingesp?*

- *wona twito loop*

---

**Repeat** (up to 8 rounds)
depends on number + animal

depends on number + animal + movement
depends on $\emptyset$ depends on number depends on number + animal
Experiment 1 results (N=94)

Proxy for native speakers
(≈ no simplification)

Proxy for non-native speakers
(simplification)

Suffix error (normalised edit distance)

Round

Complexity

Quantifier
Noun
Verb
Experiment 2: investigating the effect of population size and proportion of non-native speakers on a single step of intergenerational transmission
Experiment 1 participants

Round 2
-a -o -an, -a -op -asp, -a -o -an, ...

Round 8
-a -o -an, -ak -op -asp, -a -o -an, ...

Round 2
-ak -o -an, -ak -op -esp, -a -o -esp, ...

Round 8
-a -o -an, -ak -op -asp, -a -o -an, ...

Round 2
-a -o -esp, -ak -o -esp, -a -o -esp, ...

Round 8
-a -o -an, -ak -op -asp, -a -o -an, ...
Experiment 1 participants

Round 2
-a -o -an, -a -op -asp, -a -o -an, ...

Round 8
-a -o -an, -ak -op -asp, -a -o -an, ...

Round 2
-ak -o -an, -ak -op -esp, -a -o -esp, ...

Round 8
-a -o -an, -ak -op -asp, -a -o -an, ...

Round 2
-a -o -esp, -ak -o -esp, -a -o -esp, ...

Round 8
-a -o -an, -ak -op -asp, -a -o -an, ...

Experiment 2 participant
Experiment 1 participants

- Round 2
  - a -o -an, -a -op -asp, -a -o -an, ...

- Round 8
  - a -o -an, -ak -op -asp, -a -o -an, ...

- Round 2
  - ak -o -an, -ak -op -esp, -a -o -esp, ...

- Round 8
  - a -o -an, -ak -op -asp, -a -o -an, ...

- Round 2
  - a -o -esp, -ak -o -esp, -a -o -esp, ...

- Round 8
  - a -o -an, -ak -op -asp, -a -o -an, ...

- ...
Experiment 1 participants

Round 2
-a -o -an, -a -op -asp, -a -o -an, ...

Round 8
-a -o -an, -ak -op -asp, -a -o -an, ...

Round 2
-ak -o -an, -ak -op -esp, -a -o -esp, ...

Round 8
-a -o -an, -ak -op -asp, -a -o -an, ...

Round 2
-a -o -esp, -ak -o -esp, -a -o -esp, ...

Round 8
-a -o -an, -ak -op -asp, -a -o -an, ...

Experiment 2 participant
## Experiment 2 (N=522)

Small (2) or Large (8) population

All Native, Mixed or All Non-native input

<table>
<thead>
<tr>
<th></th>
<th>All Native</th>
<th>Mixed</th>
<th>All Non-native</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>Labels from 2 Round 8 learners</td>
<td>Labels from 1 Round 2 and 1 Round 8 learner</td>
<td>Labels from 2 Round 2 learners</td>
</tr>
<tr>
<td>Large</td>
<td>Labels from 8 Round 8 learners</td>
<td>Labels from 4 Round 2 and 4 Round 8 learners</td>
<td>Labels from 8 Round 2 learners</td>
</tr>
</tbody>
</table>
Experiment 2 results

All native > Mixed (b=-0.63, SE=0.22, p=.004)
Mixed ≈ All Non-native (b=-0.18, SE=0.22, p=.43)
No effect of population size (b=0.16, SE=0.18, p=.38)
Interaction (b=0.96, SE=0.45, p=.031)

Plot shows only round 8 data (N=222)
Stats use all N=522
Experiment 2 results

Data composition
- All Native
- Mixed
- All Non-native

All native > Mixed (b=-0.63, SE=0.22, p=.004)
Mixed ≈ All Non-native (b=-0.18, SE=0.22, p=.43)
No effect of population size (b=0.16, SE=0.18, p=.38)
Interaction (b=0.96, SE=0.45, p=.031)

Small population: Mixed-All Non-native contrast is significant
Large population: All Native-Mixed contrast is significant

Plot shows only round 8 data (N=222)
Stats use all N=522
Summary of Part 3

- Experiment 1: Learners simplify artificial language morphology (as you’d expect from parts 1-2), and early learners do so more.
- Experiment 2: There are measurable consequences of learning from non-native-like early learners even in a single generation.
- Learning from early learners / non-native speakers is a plausible mechanism driving the correlation between population size, proportion of non-native speakers, and language complexity.
- In relation to part 1: learning pressure depends on population characteristics and might therefore differ in different populations.
Conclusions

• Language is transmitted via repeated **learning** and **use**
• Language is shaped by these processes
• These processes create the regular compositional structure that makes language so powerful
• We can understand the distribution of **exceptions** to those regularities (within and across languages) in terms of those same processes