DESIGNING THE RIGHT STIMULI FOR YOUR POPULATION AND GOALS

EDINBURGH VIRTUAL WORKSHOP ON ARTIFICIAL LANGUAGE LEARNING 2021

JAMIE WHITE (UCL)

SLIDES

Will be available on my webpage:

http://www.ucl.ac.uk/~ucjtcwh/

(or google ‘james white ucl’)

BRAINSTORMING:
WHAT ARE SOME FACTORS TO CONSIDER WHEN DESIGNING STIMULI?

DESIGNING STIMULI: FACTORS TO CONSIDER

+ TIPS AND ADVICE

IMPLICIT VS EXPLICIT TASKS

Highly Implicit ⟷ Highly Explicit

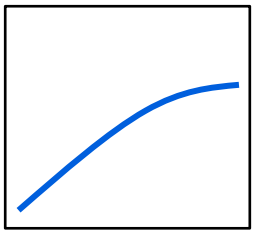
TASK SPECS

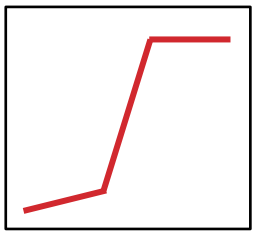
- No feedback
- Vague instructions
- Pattern difficult to describe
- Alternations not obvious/ phonotactics only

- Feedback
- Explicit instructions
- Pattern easy to describe
- Obvious alternations

OUTCOMES

- Promotes intuitive learning
- Gradual learning curve →
- More like natural learning?
- More difficult to design
- Much longer experiment





- Promotes rule seeking
- ← Sharp learning curve
- Less like natural learning
- Less difficult to design
- Quicker experiment

See Moreton & Pertsova (2017), *BUCLD*

PHONEME/PHONETIC INVENTORY

Task: Designing a vowel harmony experiment with American English speakers.

Vowel inventory:

		Front		Back	
		–round	+round	–round	+round
High	Tense	i			u
	Lax	ɪ			ʊ
Mid	Tense	e͡ɪ			o͡ʊ
	Lax	ɛ		ʌ	ɔ
Low		æ		ɑ	

PHONEME/PHONETIC INVENTORY

Use sounds that are phonemic/contrastive or easily distinguished in participants' L1.

Example: English speakers

- Palatalisation: [tusut] ⟶ [tusutʃi] ✓
- Palatal laterality: [ʎaka] ⟶ [lajaka] ✗

Non-contrastive sounds that are easily distinguished might be okay.

- English: [pataki] ⟶ [pataχ] (probably ok)
- But, there might be uncertainty about how these will be categorised featurally.

Unsure? Do a pilot study including a discrimination experiment.

WORD SHAPE AND LENGTH

Some possible stem types

CV	CVC	CVCV	CVCVC	CVCCVC	CVCVCVCV
[ka]	[kas]	[kasi]	[kasit]	[kaspit]	[kasitapi]
[ti]	[tik]	[tiku]	[tikun]	[tikmun]	[tikunali]
[bu]	[bus]	[busa]	[busak]	[bustak]	[busakapu]
[si]	[sin]	[sini]	[sinip]	[sinkip]	[sinipika]
[la]	[laf]	[lafa]	[lafan]	[laftan]	[lafanubi]
[gu]	[gut]	[guti]	[gutil]	[gutsil]	[gutiluna]

WORD SHAPE AND LENGTH

CV syllables are good as a default, general-use template.

Word length:

- **CV(C)** – probably too short
  - Difficult to avoid sounding like real words.
  - Difficult to have many distinct stems or sufficient variety.
- **CVCVCVCV** – probably too long (unless needed by design)
  - May start to feel overwhelming to participants; may obscure pattern of interest.
  - More difficult to control everything.
- **CVCV(C)** – happy medium
  - Good word size for most uses.

CONTROLLING EXTRANEOUS VARIABLES

[kasat] ⟶ [kasatʃi]

[kupit] ⟶ [kupitʃi]

[lasut] ⟶ [lasutʃi]

[kinat] ⟶ [kinatʃi]

[kulad] ⟶ [kuladʒi]

[kanid] ⟶ [kanedʒi]

[kisud] ⟶ [kisudʒi]

[kubid] ⟶ [kubidʒi]

[lanap] ⟶ [kasapi]

[pukip] ⟶ [pukipi]

[batup] ⟶ [batupi]

[nilap] ⟶ [nilapi]

[punak] ⟶ [punaki]

[tapik] ⟶ [tapiki]

[nituk] ⟶ [nituki]

[nusik] ⟶ [nusiki]

CONTROLLING EXTRANEOUS VARIABLES

Avoid accidental generalisations in the items.

- Use consonants and vowels about equally, in total and in each position (other than as required by the design).
- Especially important if there are few items. With more items, roughly balanced is usually fine.
- Depending on the design and number of items, it may even be desirable to balance CV syllables and C…C / V…V sequences.

Avoid items that sound like real words/phrases

- English: [toni], [gazibo], [da'ki], [gosi]
- French: [leba] (anything starting with [le], [la], [me], [ma], …)

CONTROLLING EXTRANEOUS VARIABLES

Inventory size

- Small phoneme inventory:
  - Easier to control things, but also much more important to control things.
  - Words start sounding really similar to each other (which could be good or bad).

- Larger phoneme inventory:
  - Harder to control perfectly, but ‘roughly balanced’ is often enough.
  - Words sounds more distinct; more realistic sounding?

CONTROLLING EXTRANEOUS VARIABLES

Stress

- Which syllable? [ˈpikuta] -- [piˈkuta] -- [pikuˈta]

- Be mindful that stress may impact interpretation and/or salience of various parts of the word with speakers of a stress language.

- Does the stress shift with morphemes added?

[ˈpikut] ⟶ [ˈpikutʃi] [piˈkut] ⟶ [piˈkutʃi]

[ˈpikut] ⟶ [piˈkutʃi] [piˈkut] ⟶ [pikuˈtʃi]

- I tend to favour a stem-bounded stress pattern (no stress shift).
  - Probably less common typologically (though attested), but reduces extraneous changes in the stimuli if stress is not a variable of interest.

HOW MANY ITEMS/TRIALS?

You need enough items/trials to:

- Allow participants to learn pattern(s).
- Discourage people from focusing too much on memorising specific items.
- Ensure that there are not accidental patterns for people to find / focus on.

Too many items/trials can:

- Cause a ceiling effect, if not using the Extrapolation/Poverty of the Stimulus paradigm.
- Cause the experiment to be long and sap motivation late in the study.

MANY DISTINCT ITEMS VS. REPEATED ITEMS

Many distinct items:

- Encourages finding broad generalisations. (Participants cannot remember individual items well).
- OK for properties of words (sounds/syllable types) to be ‘roughly balanced’.
- Harder to set up (e.g., recording time, balancing).

Fewer items / more repetition:

- More likely that participants remember/focus on specific items.
- Properties of words should be very carefully controlled.
- Easier to set up.

TRAINED VS. NOVEL ITEMS

Trained/Old/Familiar items:

- Items from training repeated in the test phase.

Novel items:

- Items only seen in the test phase.

Depending on the design, it is often a good idea to include both types in the test phase.

- **Trained items:** can let you ensure participants actually learned something and paid attention.
- **Novel items:** show you whether participants extracted a general pattern (because Trained items could be memorised/recognised).

CRITICAL & FILLER ITEMS

Critical items: items needed to train and/or test the pattern of interest.

Filler items: items not important for analysis, but added to round out the experiment. They are included to:

- Distract from the patterns/sounds of interest.
- Establish phonotactic facts (or other properties) about the language.
- Make the language seem more realistic, with a variety of words.

RECORDED VS SYNTHESISED

Recorded

Synthesised

RECORDED VS SYNTHESISED

Recorded

Pros:

- Sounds more natural.
- Can ask speaker to speak in a certain way.

Cons:

- More variability (may be less of a concern with many words).
- Time-consuming (especially with many words).

Synthesised

Pros:

- May sound unnatural.
- Cannot ask speaker to add specific features, etc.

Cons:

- Less variability.
- Quicker/easier (especially with many words).

MORE RECORDING ISSUES

Who will record it?

- Native speaker of participants' L1 or a different language?

How sounds are produced?

- Stops with aspiration?
- Monophthongs or diphthongs?
- Stress pattern?

Control/measure/manipulate any properties?

- Equalising for intensity is standard.
- Others (syllable duration, f0, etc.) may be desirable depending on design.

PICTURES OR NO?



Using pictures

- Makes task more interesting.
- May distract (somewhat) from pattern of interest.
- Some participants will look for semantic patterns (e.g., use [-fi] when the picture is animate, round, a fruit, red, …)
- Must be very careful to ensure that there are no accidental properties of the pictures confounded with the pattern of interest!

Not using pictures

- Prevents participants from (accidentally) looking for semantic patterns.
- Feels a bit boring – motivation may sag fairly quickly!

EMPHASISE OR OBSCURE THE PATTERN OF INTEREST?

Ways to emphasise the pattern:

- Explicit task (explicit alternations, feedback, etc.)
- Shorter words
- Fewer filler items
- Put the target sounds in prominent/salient positions
- No pictures

Ways to obscure the pattern:

- Implicit task (no explicit alternations, no feedback, etc.)
- Longer words
- More filler items
- Add a ‘red herring’ pattern and/or a ‘red herring’ task in instructions
- Include pictures

QUESTIONS?

REFERENCES

- Moreton, Elliott, and Katya Pertsova (2017). Implicit and explicit processes in phonotactic learning. In: Jennifer Scott and Deb Waugtal (eds.), *Proceedings of the 40th Boston University Conference on Language Acquisition* (BUCLD 40), pp. 277--290.