Tutorial in Experimental Syntax

Steven Foley (University of Southern California) & Léa Nash (University of Paris 8) contact: foleys@usc.edu & lea.nash@univ-paris8.fr download slides here: stevenrfoley.github.io SCCC-3 • Ilia State University • Tbilisi, Georgia • October 2–4, 2023

1. Introduction

Contrasting methodologies in **formal** linguistics vs. **experimental** linguistics

- **Cooking** accumulating evidence via introspection / fieldwork / corpus research
- **Baking** collecting quantifiable behavior in batches in controlled conditions

Formal syntax vs. Experimental syntax vs. Sentence processing

- All concerned with theories of grammar & linguistic knowledge
- Is competence or performance more central? Is **time** important, or memory?
- Documentary / descriptive goals vs. Theoretical ones
- Subtle contextual factors? Subtle acceptability gradience?

1. Introduction

Goals of this tutorial

Introduce concepts and walk through a case study

• Motivate and design an acceptability judgement study on Georgian number agreement

Give an answer to: "Where do I start with experimental work?" – **How is it useful?**

- Keep it **simple** Compare minimal pairs or quartets
- Have a **hypothesis** How and why will a grammatical factor influence measured behavior?

Outline

- 1. Introduction
- 2. Core Ideas
- 3. Experimental Design
- 4. Case study
- 5. Q&A / Discussion

2. Core Ideas

- 2.1 Syntax
- 2.2 Psycholinguistics
- 2.3 Linking Hypotheses

2.1 Syntax

Research questions in syntactic theory

What factors influence number agreement in Georgian? In languages generally?

- Grammatical relation, linear order, animacy, discourse status, etc.
- (1) მასწავლებლ**-ებმა** ბავშვი **დამალეს** / ***დამალა** teacher-**PL:ERG** child:NOM hide:TR:AOR:**3PL.A** / ***...3SG.A** "The teachers hid the child"
- (2) მასწავლებლ**-ებს** ბავშვი **დაემალა** / **?დაემალათ** teacher**-PL:DAT** child:NOM hide:NACT:AOR:**3SG.S**>3.IO / **?...3PL.IO** "The child hid from the teachers"

2.2 Psycholinguistics



Research questions in psycholinguistic theory

How is the acceptability of number agreement similar to (e.g.) color perception?

- What are the properties of Agreement Attraction Illusions?
- (3) a. [_{SG} The key to [_{SG} the cabinet]] is on the table. /...*are...
 b. [_{SG} The key to [_{PL} the cabinets]] is on the table. /...?are...
 c. [_{PL} The keys to [_{SG} the cabinet]] are on the table. /...?is...
 d. [_{PL} The keys to [_{PL} the cabinets]] are on the table. /...*is...

2.2 Psycholinguistics



Research questions in psycholinguistic theory

How is the acceptability of number agreement similar to color perception?

- What are the properties of Agreement Attraction Illusions?
- (4) a. [_{SG} Which flower] has [_{SG} the gardener] watered already? /...*have...
 b. [_{PL} Which flowers] ?has [_{SG} the gardener] watered already? /...?have...
 c. [_{SG} Which flower] have [_{SG} the gardeners] watered already? /...?has...
 d. [_{PL} Which flowers] have [_{SG} the gardeners] watered already? /...*has...

2.3 Linking hypotheses

Theories of behavioral measurements

Which grammatical factors could be manipulated in your experiment?

- Subject vs. IO agreement; SG vs PL controller; SG vs PL attractor; Overt vs. *pro*; Discourse status How will participants' perception of those factors influence measured behavior?
- "They will be more sensitive to agreement errors involving this factor because..."
- More sensitive to a grammatical factor = Correct uses rated high & Incorrect uses rated low

3. Experimental Design

- 3.1 Practical Limitations
- 3.2 Planning Stimuli
- 3.3 Narrowing the Hypothesis

3.1 Practical Limitations

Designing an experiment is a numbers game, a balancing act

- **Budget:** How much do I have to pay participants?
- **Participants:** How many participants can I recruit, and from what populations?
- Length: How many observations per participant do I need?
- **Task:** What kind of behavioral measure will be most informative?

3.1 Practical Limitations

Who are my participants, and how are they participating?

- What social demographics are important? (Age, gender, fluency, dialect...)
- Mode of participation: in-person vs. remote/internet-based experiments How much should I pay them?
- Depends on the length of experiment, difficulty of task, and goodwill of participants
- What will **incentivize and fairly compensate** careful attention?

3.1 Practical Limitations

Concerns relevant to choosing between experimental methods

• How subtle is the contrast? Do stimuli need to be auditory? Is time an important factor?

Off-line measures

• Untimed judgements, given upon reflection

Stimulus:

What did Napoleon have invented?

Prompt:

7 - 6 - 5 - 4 - 3 - 2 - 1 Very good Very bad

 On-line / real-time measures E.g., word-by-word RTs – incremental processing 				
Stimulu <i>What</i>	s:			
	did			
		Napoleon		
			have	
				invented?

How many observations (experimental trials) per stimulus type do I need?

• What kinds of **filler trials** would be useful, and how many? **Comprehension** checks?

Trial 1: Target Stimulus, Version A The key to the cabinets is on the table.

Trial 2: Filler Napoleon had glass jars invented.

Trial 3: Target Stimulus, Version B The film about the glaciers were very interesting.

Trial 4: Filler The horse raced past the barn fell.

. . .

Trial N: Target Stimulus, Version A *The letter from the spies has been burned.*



Studying structures, not particular sentences/words

• Lists and Itemsets – lowering noise associated with particular lexicalizations

Condition A (Gram. SG-Agr)

- **Itemset 1** The key to the cabinets <u>is</u> on the table.
- **Itemset 2** The film about the glaciers was very interesting.
- **Itemset 3** The letter from the spies <u>has</u> been burned.

•••

Itemset N The chemical in the cookie <u>is</u> totally harmless.

Condition B (Attracted PL-Agr) ...

The key to the cabinets \underline{are} on the table

The film about the glaciers \underline{were} very interesting.

The letter from the spies have been burned.

The chemical in the cookies <u>are</u> totally harmless.

Purpose of non-target trials (fillers)

- **Distraction:** Preventing participants from developing a sneaky strategy
- Scale-setting: Providing context to anchor judgements / expectations
- **Piloting:** Mini-experiments (unpublishable ones) that test hypotheses for future studies

Ideally, a mix of structures similar to and different from the targets



Rules of thumb for experiment size Bare minimum

- 4 observations per stimulus type; 20 participants
- Standard for psycholinguistics
- 8 observations per stimulus type; 60 participants
- Pulling out all the stops
- 12 observations per stimulus type; 300 participants

3.3 Narrowing the Hypothesis

Experiments are good for **testing hypotheses**, not generating them

- Test theoretically motivated questions
- Build on anecdotal observation (fieldwork, corpus patterns)
- Don't bother verifying very crisp judgements (except as baselines / in fillers)
- Focus on a few grammatical factors that can create a **minimal pair/tuplet**
- Experiments are good at isolating factors and their interactions

4. Case Study

- 4.1 Estimating Experiment Size
- 4.2 Choosing Design Factors
- 4.3 Choosing Experimental Task
- 4.4 Planning Stimuli

4.1 Estimating Experiment Size

Goals and Constraints

Practical limitations allows 40 participants at 96 trials each (< 30 minutes)

- $96 \div 3 = 32$ critical trials
- 32 itemsets = 4 observations per participant of 8 conditions, or 8 observations of 4 conditions
- 96 32 = 64 filler trials

4.2 Choosing Design Factors

What grammatical factors could be manipulated in principle?

- **Type** of agreement (3SG or 3PL)
- **Grammatical role** of PL argument (Tr Subj, Intr Subj, Dir Obj, Ind Obj...)
- Discourse status of PL argument (Given/Topic vs. New/Focus)
- Coargument features (3SG, 3PL, 1ST/2ND)
- Word Order (SOV, SVO...)
- **Animacy** of PL argument (Animate, Inanimate)
- **Aspect** of verb (Imperfective, Perfective)



4.2 Choosing Design Factors

Illustrating possible grammatical manipulations

- A=3PL ბავშვებმა ექიმი დამალეს child:**PL:ERG** doctor:NOM hid:TR:**3PL.A** "The children hid the doctor"
- P=3PL ბავშვმა ექიმები *დამალათ child:ERG doctor:**PL:NOM** hid:TR:**3PL.P** "The child hid the doctors"
- S=3PL ბავშვები ექიმს დაემალნენ / * child:**PL:NOM** doctor:DAT hid_from:**3PL.S** "The children hid from the doctor"
- G=3PL ბავშვი ექიმებს ?დაემალათ / ?დაემალა child:NOM doctor:**PL:DAT** hid_from:**3PL.G** hid_from:TR:**3SG.G** "The child hid from the doctors"

დამალა hid:TR:**3SG.P**

SG Agr

*დამალა

hid:TR:3SG.A

***დაემალა** hid_from:**3SG.S** The S/G contrast is most is most

4.2 Choosing Design Factors

Illustrating possible contextual manipulations

- N1=TOP გუშინ ვნახე ბავშვ(ებ)ი ეზოში yesterday saw:1SG.A child(:PL):NOM garden:in "Yesterday I saw **a child / children** in the garden"
- S=3PL ბავშვები ექიმს დაემალნენ / child:PL:NOM doctor:DAT hid_from:3PL.S "The children hid from the doctor"
- G=3PL ბავშვი ექიმებს ?დაემალათ / child:NOM doctor:PL:DAT hid_from:3PL.G "The child hid from the doctors"

?დაემალა hid_from:TR:3SG.G

hid_from:3SG.S

*დაემალა

4.3 Choosing Experimental Task

Option 1: Separate judgements for each agreement type

Trial 1

Context:

გუშინ ვნახე ექიმები ეზოში. "Yesterday I saw doctors in the garden"

Stimulus:

ბავშვი ექიმებს დაემადა. "A child hid [SG.Agr] from the doctors."

Prompt:

7 - 6 - 5 - 4 - 3 - 2 - 1ปงถางธ์ งงตลกปงถางธ์ เขาอก"very good""very bad"

Trial 2

Context:

გუშინ ვნახე ექიმები ეზოში. "Yesterday I saw doctors in the garden"

Stimulus:

ბავშვი ექიმებს დაემადათ. "A child hid [PL.Agr] from the doctors."

Prompt:

7 - 6 - 5 - 4 - 3 - 2 - 1 dsqnsb зsრдп dsqnsb gyqn "very good" "very bad"

4.3 Choosing Experimental Task

Option 2: Separate judgements for each agreement type

Trial 1

Context:

გუშინ ვნახე ექიმები ე8ოში. "Yesterday I saw doctors in the garden"

Stimulus:

Prompt:

. *"დაემაღა" სჯობს "დაემაღათ" სჯობს* SG.Agr is better PL.Agr is better

Settling on a design

- Factor 1: Grammatical role of PL argument = {Nonact Subj, Ind Obj}
- Factor 2: Context = {Subj Topic, IO Topic}
- (Factor 3: Agreement type = {3SG, 3PL})

Task 1: 8 conditions per itemset, 32 itemsets = 4 observations per condition Task 2: 4 conditions per itemset, 32 itemsets = 8 observations per condition

What verbs to choose?

- Good place to start: Georgian National Corpora
- (What kinds of IOs? Obliques, benefactors, possessors, reciprocals...)

Gippert et al. 2012

GRC :: Word list



5. Questions

References

Dillon, Brian, Adrian Staub, Joshua Levy, and Charles Clifton. 2017. Which noun phrases is the verb supposed to agree with?: Object agreement in American English. Language, 93(1). 2001. • Gippert, Jost, Paul Meurer, and Manana Tandashvili. 2012. The Georgian National Corpus Project. • Tuite, Kevin. 1998. Kartvelian morphosyntax: Number agreement and morphosyntactic orientation in South Caucasian languages. Dissertation, University of Montreal. • Wagers, Matt, Ellen Lau, and Collin Phillips. Agreement attraction in comprehension: Representations and processes. Journal of Memory and Language, 61.