

South Caucasian Workshop
The U Chicago Center Paris

Tutorial on Experimental work

Maria Polinsky & Stavros Skopeteas

Part II. experimental studies on Georgian

by Stavros Skopeteas

Paris, September 23, 2016

experimental methods in grammatical research

linguistic research based on two sources of evidence:

- observational research: possible linguistic entities in corpora
- introspective research: grammaticality judgments

recent focus on repeated-observation designs:

- quantitative corpus studies or production experiments
- questionnaire studies on speakers' intuitions

☞ **tutorial, part I:** offline methods, not informative for cognitive processes, crucial for testing grammatical hypotheses

motivation for these developments: increase reliability, estimate gradience

contents of this part

experimental data in language comparison

- increase reliability

illustrative study: focus in OV languages (Georgian, Armenian, Turkish)
(questionnaire study)

speaking in the lab vs. speaking in the world

- same contrasts, different sources of variation

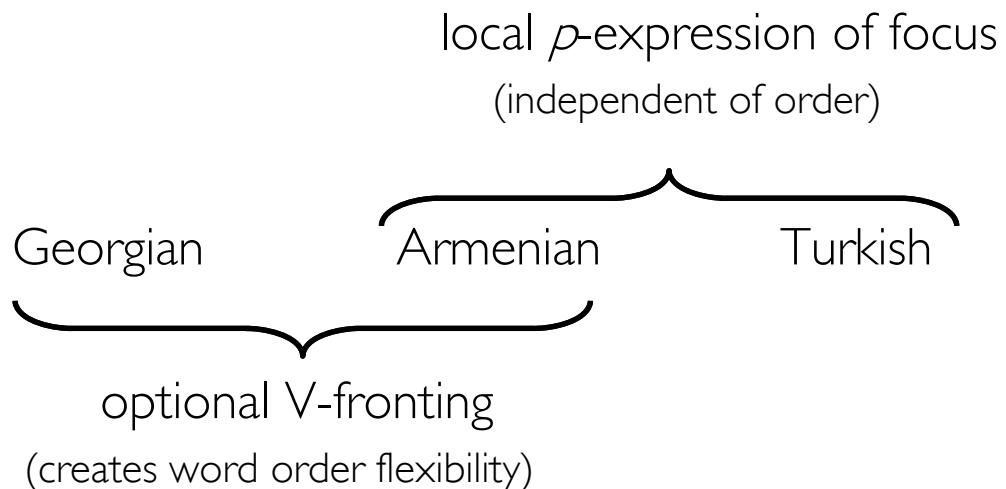
illustrative study: prosodic effects of order and focus in Georgian
(scripted speech vs. corpus data)

experimental data in language comparison

grammatical background

Georgian, Armenian and Turkish

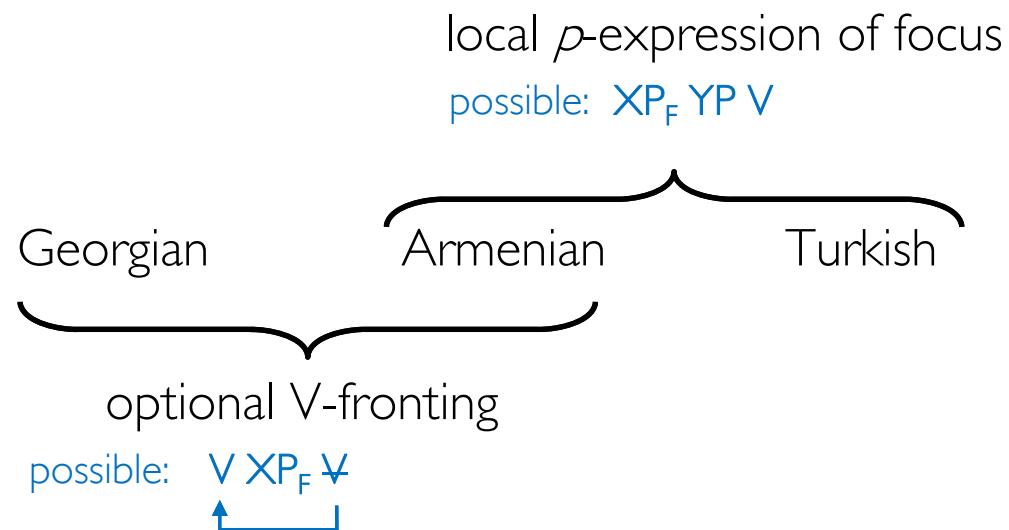
are OV languages, but differ in details:



grammatical background

Georgian, Armenian and Turkish

OV languages, differing in details:



method

illustration

Q	զո՞ն	օպո՞զա	ծյուրտօ?
	who	found	ball?
A	քիրամ	ծյուրտօ	օպո՞զա.
	Kira	ball	found

(does not fit) 1 2 3 4 5 6 7 (fits)

factorial design

CONTEXT (question):

Sentence-focus|Subject-focus|Object-focus|Verb-focus

WORD ORDER (answer):

SVO|SOV|OVS|OSV|VSO|VOS

random factors

48 lexicalizations, 48 speakers

result

Georgian

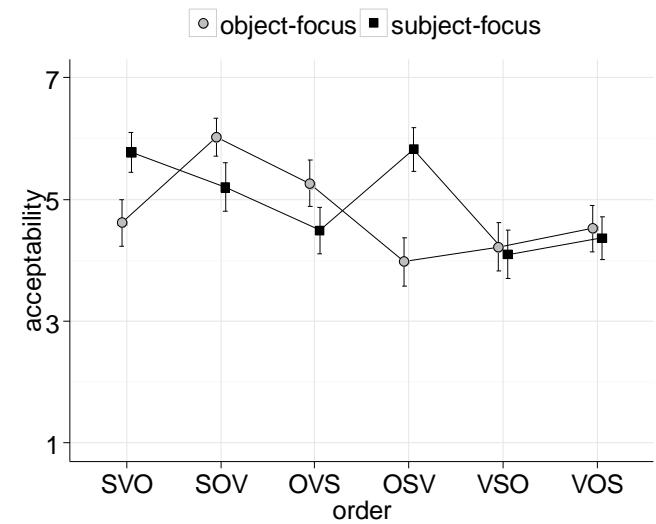
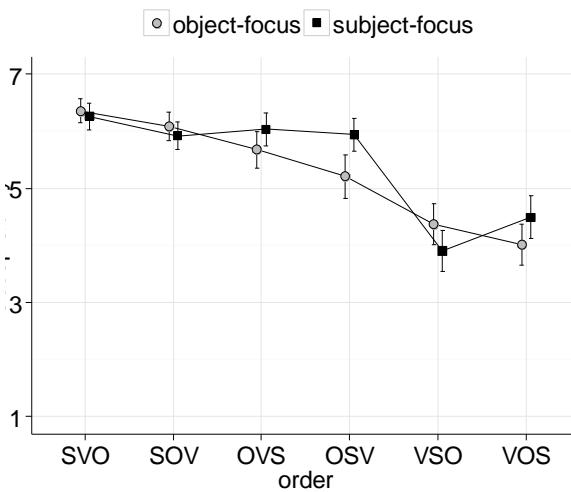
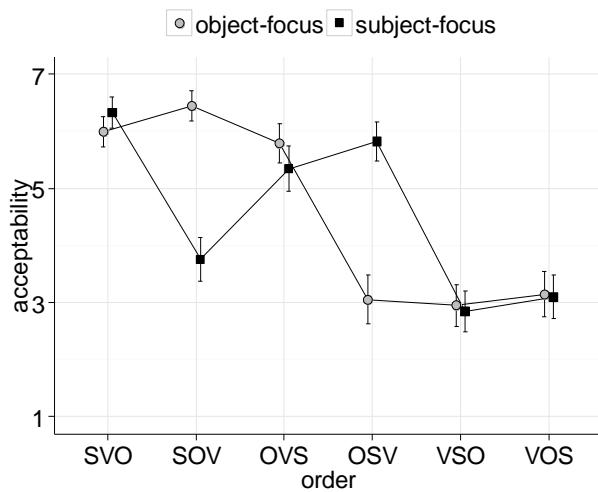
S_F	$X_F V$	$V X_F$	$X_F V$
O_F	$V X_F$	$X_F V$	$X_F V$
	SVO	SOV	OVS

Armenian

S_F	$X_F V$	$X_F YV$	$V X_F$	$X_F V$
O_F	$V X_F$	$X_F V$	$X_F V$	
	SVO	SOV	OVS	OSV

Turkish

S_F	$X_F V$	$X_F YV$	$X_F V$
O_F		$X_F V$	$X_F V$
	SVO	SOV	OVS



discussion

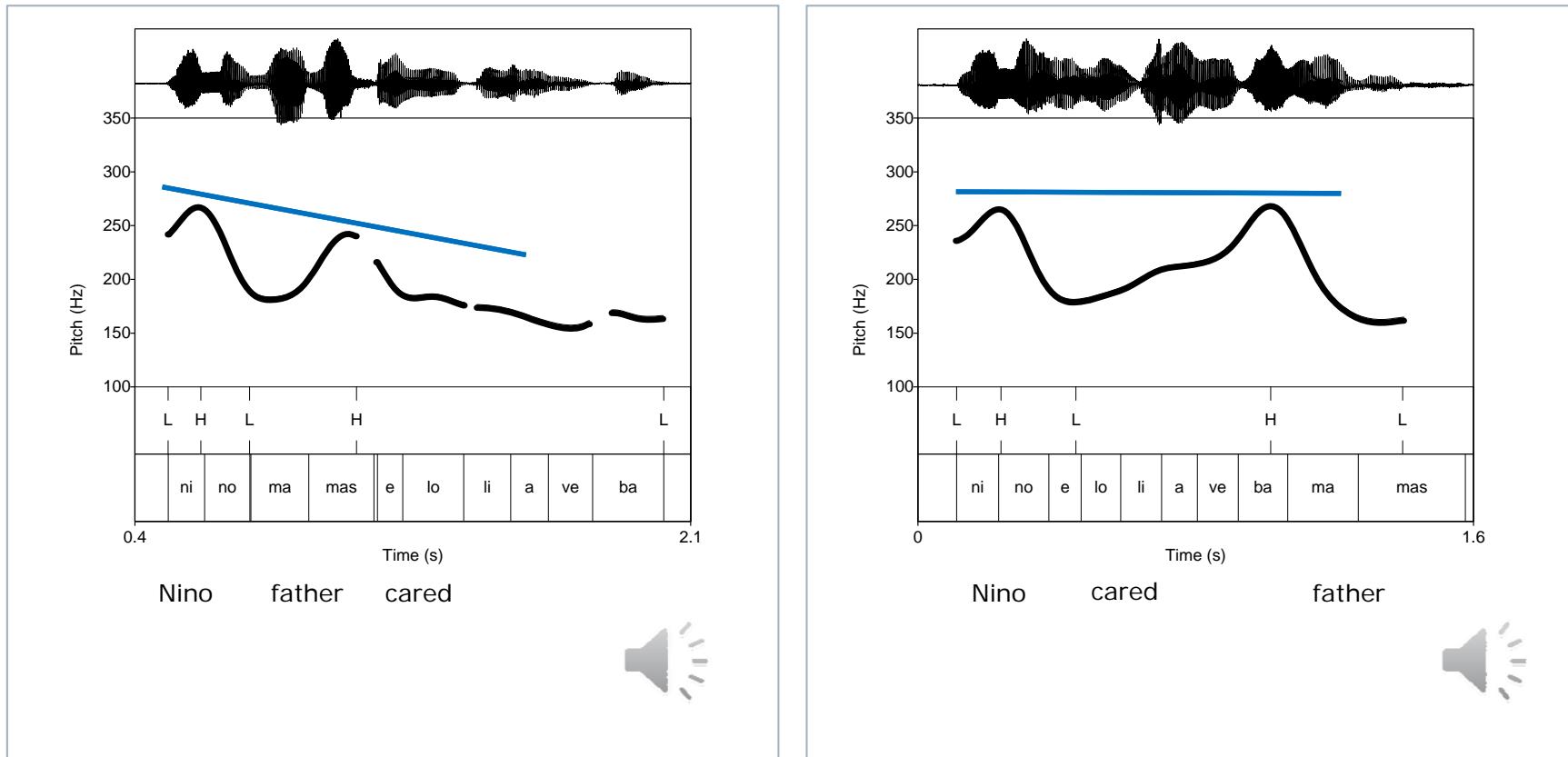
on the reliability of controlling sources of variance in a Between Languages design

- ☞ tutorial, part I: acceptability vs corpus: what do we learn about modularity from the mapping between different data types?

speaking in the lab vs. speaking in the world

observation

H-phrase tones are integrated within a single domain of downstep in SOV but not in SVO.



method

Context * Order

(all and only felicitous permutations)

	SOV	SVO	OSV	OVS
all_F	✓	✓		
VP_F	✓	✓		✓
S_F		✓	✓	✓
V_F	✓	✓		
O_F	✓	✓		✓

Items (4)

nino mama-s e-loliav-eb-a.
Nino(NOM) father-DAT O3-care-THM-DEP.S3.SG
Nino cares about the father.

Context questions

ALL: What happens?

VP: What do we hear about Nino?

S: Who cares about the father?

eight speakers

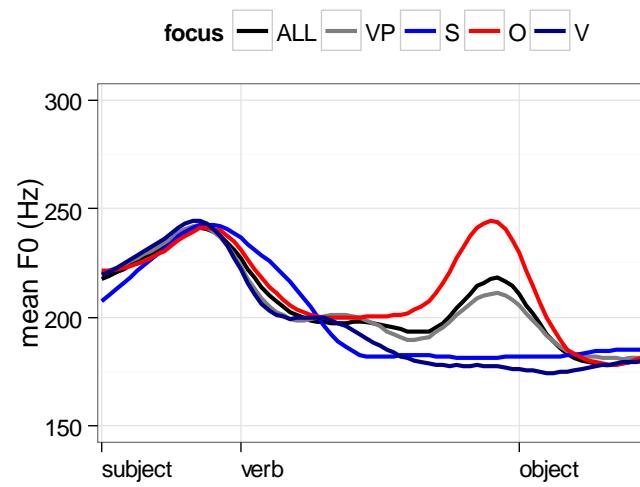
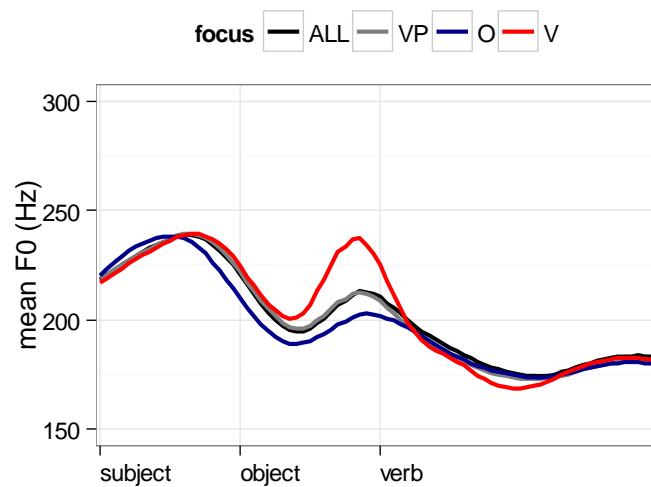
two repetitions per item

Total= 832 tokens

results

two sources of non-downstepped H-phrase tones:

- right-edge of a **non-final V**
- preceding a **focused XP**



theoretical relevance

reported phrase-languages

ALIGN(focus-*i*;φ-*j*)

Align the *i*-edge of the **focus** with the *i*-edge of a prosodic phrase.

(Büring 2010, Féry 2013)

hypothesis

In a subset of these languages (among else Georgian), the prosodic edges are mapped on nucleus (not on focus)

ALIGN(nucleus-*i*;φ-*j*)

Align the *i*-edge of the **nucleus** with the *i*-edge of a prosodic phrase.

syntax-phonology mapping

stress-based language

maximally embedded: stressed

*		*
*		*
*	*	*
<hr/>		
[S	[V	[O]]
*		
*	*	
*	*	*
<hr/>		
[S	[[O]	V]

phrase-based language

maximally embedded: separated

()	
(()
<hr/>		
[S	[V	[O]]
()	
(()
<hr/>		
[S	[[O]	V]

speaking in the lab vs. speaking in the world

external validity?

- is the inference from the sample to the world valid
- or is it an artefact of the lab situation?

background:

- current discussion in phonetics about scripted data.
(Xu 2010; Wagner, Trouvain, & Zimmerer, 2014, among else)

Georgian corpus

material

Narratives

Activity description

Ancestor story

Path description

Event description

Comparative description

Style

non-scripted data; not spontaneous, but less attention to speech; naturalistic behavior (laughter, spontaneous speech planning)

speakers

24 speakers

- residents of Tbilisi;
- speaking Georgian as a first language;
- generally high education level;
- age range: 20-58 (average 29.3).

corpus annotation

phenomena of interest

- order: XP V YP vs. XP YP V
- focus: approximated through givenness

phenomena outside the scope of the intended generalization => excluded

- non-declaratives, embedded clauses
- laughter, disfluencies, etc.

phenomena for which the generalization is intended to apply (random factors)

- speakers
- syllable structures
- syntactic category of the XP

speaker	XP V YP	XP YP V	dur (sec)
01	8	5	207
02	5	3	497
03	17	5	268
04	11	8	309
05	6	4	228
06	9	4	179
09	10	7	320
10	14	5	365
11	8	5	218
12	9	2	173
14	5	3	196
15	7	4	290
16	4	6	296
17	10	10	389
18	12	8	361
20	10	2	197
21	6	3	215
22	11	3	483
23	13	2	167
24	7	4	99
total	182	93	91 min.

results

word order effect

final vs. non-final V

measurements

voiced part of the first and last syllable

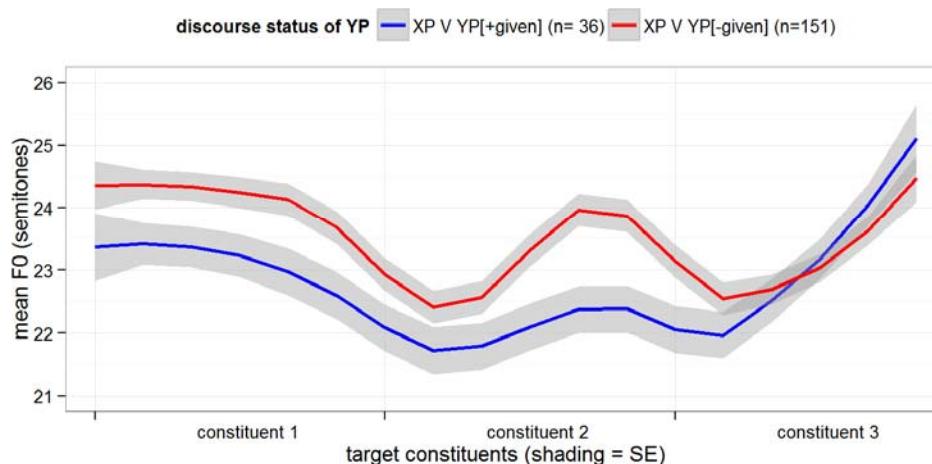
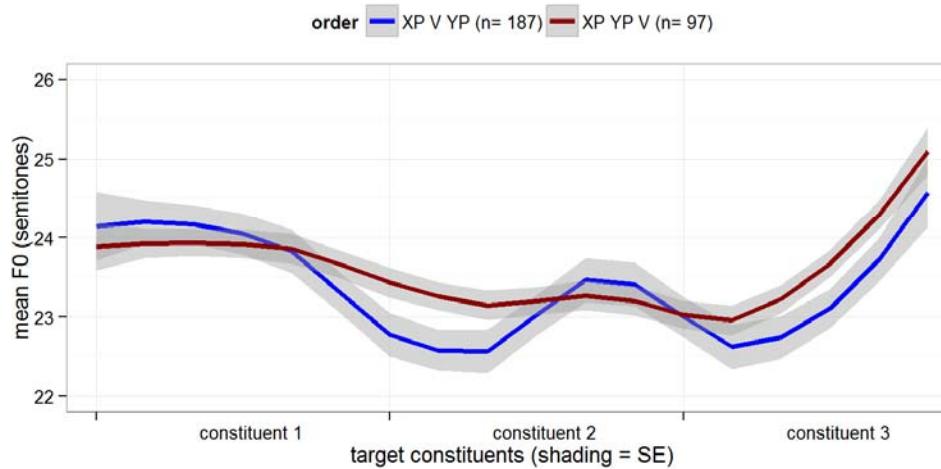
of the target constituent

F_0 -mean of three equal intervals per target σ

givenness

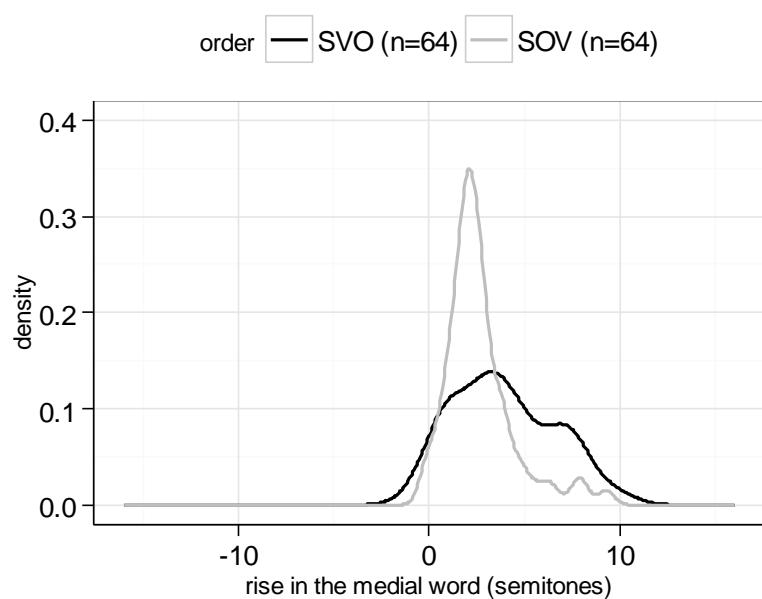
as predictor of focus

given = referent mentioned in the pre-text

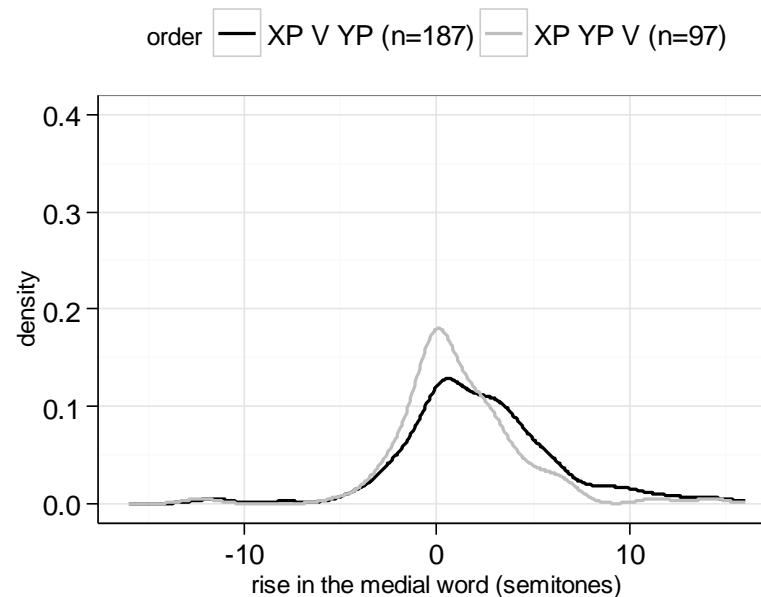


comparison between data types

scripted speech



unscripted speech



rise in the medial word:

$F_0\text{-max}$ (last syllable) – $F_0\text{-min}$ (first syllable) of the medial constituent

final discussion

- on the external validity of experimental data
 - controlling variation vs outbalancing variation
 - are there lab grammars?
- ☞ Continuing the discussion in **tutorial, part I**: limits of experimental designs, when should I come back to the traditional methods?

South Caucasian Workshop
The U Chicago Center Paris

Tutorial on Experimental work

Maria Polinsky & Stavros Skopeteas

Part II. Georgian

by Stavros Skopeteas

Paris, September 23, 2016