

# The role of 'chunks' in second language learning and teaching'

Florence Myles

York, 8<sup>th</sup> September 2016

# Outline

- Formulaic sequences/chunks: what are they?
  - In native language use
  - In second language learners
- Two studies
  - Beginners
  - Advanced learners
- Role of FS
  - In beginners
  - In advanced learners
- Implications for teaching

# Formulaic sequences: what are they?

- Speaker external definition
  - Sequence commonly found together in native speaker corpora
    - E.g. *will you marry me?* rather than *would you like to get married with me?* *Nice day today* rather than *the weather is good today*
  - Sequence which is a single semantic unit which cannot be generated from its individual constituents
    - E.g. *kick the bucket; by and large; thanks for having me*
- Speaker internal definition
  - Processing shortcut for that particular individual
  - What is formulaic in one individual is not necessarily formulaic in corpora or in the target language, and vice versa, e.g.
  - In the studies reported here psycholinguistic definition (Wray 2002; Myles et al. 1998)

# Most commonly used psycholinguistic definition

“a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar.”

Wray, 2002: p. 9

In the context of L2 learners:

- What is processed as a single unit, for two distinct reasons:
  - because the multiword unit is stored as a whole in the learner’s lexicon (it has been learnt that way)
  - because it is highly automatised (through repeated use)

# Identification issues

- How can you tell a learner has processed a sequence as one unit or not?
  - Relatively easy in beginners: FS are very different from the rest of a learner's productions:
    - Greater complexity, grammatically more advanced
    - More fluent
    - Often inappropriate (semantically, syntactically etc.)
    - Used in same form
  - Much more difficult in advanced learners
    - Evidence of a processing advantage

# Identification in beginners – example of one criterion

Greater **complexity** of sequence:

(1) *quel âge as-tu?*

what age have you? (*wh*-fronting and verb-subject inversion) = how old are you?

(Intended meaning: “how old is your brother?”)

(2) *\*Il âge frère?*

he age brother?

(Intended meaning: “how old is his brother?”)

Pupil 02, R4; produced by the same learner during the same elicitation task)

(all examples taken from study 1)

# The studies

- Early learners study datasets
  - Beginners
    - Longitudinal over 2 years; 60 learners (analysis here on a subset of 16 learners); years 7, 8, 9 (11-14 years old); up to 13 oral tasks per learner (all one-to-one with researcher or child-child pair work)
  - Post-beginners
    - Cross-sectional; 20 in each of years 9, 10, 11 (14-16 years old); 4 oral tasks (1:1 with researcher; 3 repeated from beginners study)
  - Details on [www.flloc.soton.ac.uk](http://www.flloc.soton.ac.uk)

# Advanced learners' study (Cordier 2013)

- 5 English university learners of French (similar educational background and level in French)
- Longitudinal study pre/post year abroad
- 2 data collection points:
  - End of 2<sup>nd</sup> year at university (immediately prior to compulsory year abroad in France)
  - During 7<sup>th</sup> month of stay in France
- 5 oral tasks repeated at time 1 and 2 (10 months apart): general 1:1 interview with researcher; picture based story telling task; 3 discussion tasks in pairs (approx 1 hr of data per learner each time)



# Use of FS - Beginners

- The role of FS in the development of interrogative constructions
- All complex structures involving wh-fronting and subject verb inversion, well beyond the syntactic means of beginner learners
- Interrogative chunks tracked
  - *Où habites-tu?*  
where live you (where do you live?)
  - *Comment t'appelles-tu?*  
how yourself call you (what's your name?)
  - *Quel âge as-tu?*  
what age have you (how old are you?)
  - *Quelle est la date de ton anniversaire?*  
what is the date of your birthday

# Research questions

- how do these chunks, all present in all learners at the beginning of data collection, develop over time?
- How do interrogative constructions when learners have no chunks available develop over time?
- What is the link between the two?
  
- Asking 'what's his/her name?'
- Chunk: *comment t'appelles-tu?* 2<sup>nd</sup> person reference (what's your name)
- How do learners ask 'what's his/her name?'

# 5 developmental stages for 3<sup>rd</sup> person naming questions

1. Chunk over-extended: *\*comment t'appelles-tu?* (e.g. pointing at the person on the photo)
2. Chunk over-extended, but lexical NP tagged on: *\*comment t'appelles-tu le garçon?*
3. Chunk starting to break down: *\*comment t'appelle (la fille)?*
4. Further breaking down: *comment s'appelle (un garçon)?*
5. 3<sup>rd</sup> person pronoun used: *comment s'appelle-t-il?*

# Development of interrogatives when no chunk is available

## 1. Verbless stage

*Je grand maison?* (I big house)

## 2. Non-finite verb stage

*La mère regarder le magasin?* (the mother look – non-finite – the shop)

## 3. Finite verb stage

*La mère regarde livre?* (the mother looks book)

This is in keeping with the development of the productive grammar generally

## Relationship between chunks and generative grammar

At the same point in time, we find within individual learners both:

- Highly complex chunks, delivered relatively fluidly and accurately, e.g. *comment t'appelles-tu?*
- Very simple syntax outside chunks, e.g. *euh nom?* (pointing at picture)

# Role of FS in beginners

- All learners use FS, but some are much better at memorising them;
- The learners who have the largest store of FS keep using them (i.e. they do not discard them), but they are also the learners who are most active at 'working' on them to use their constituent parts elsewhere (all 'top' learners had a large store of FS);
- FS give learners an entry into communication, before their interlanguage grammar can produce complex structures;
- FS provide a set of language samples which act as models for constructing their grammar

# The use of FS in advanced learners

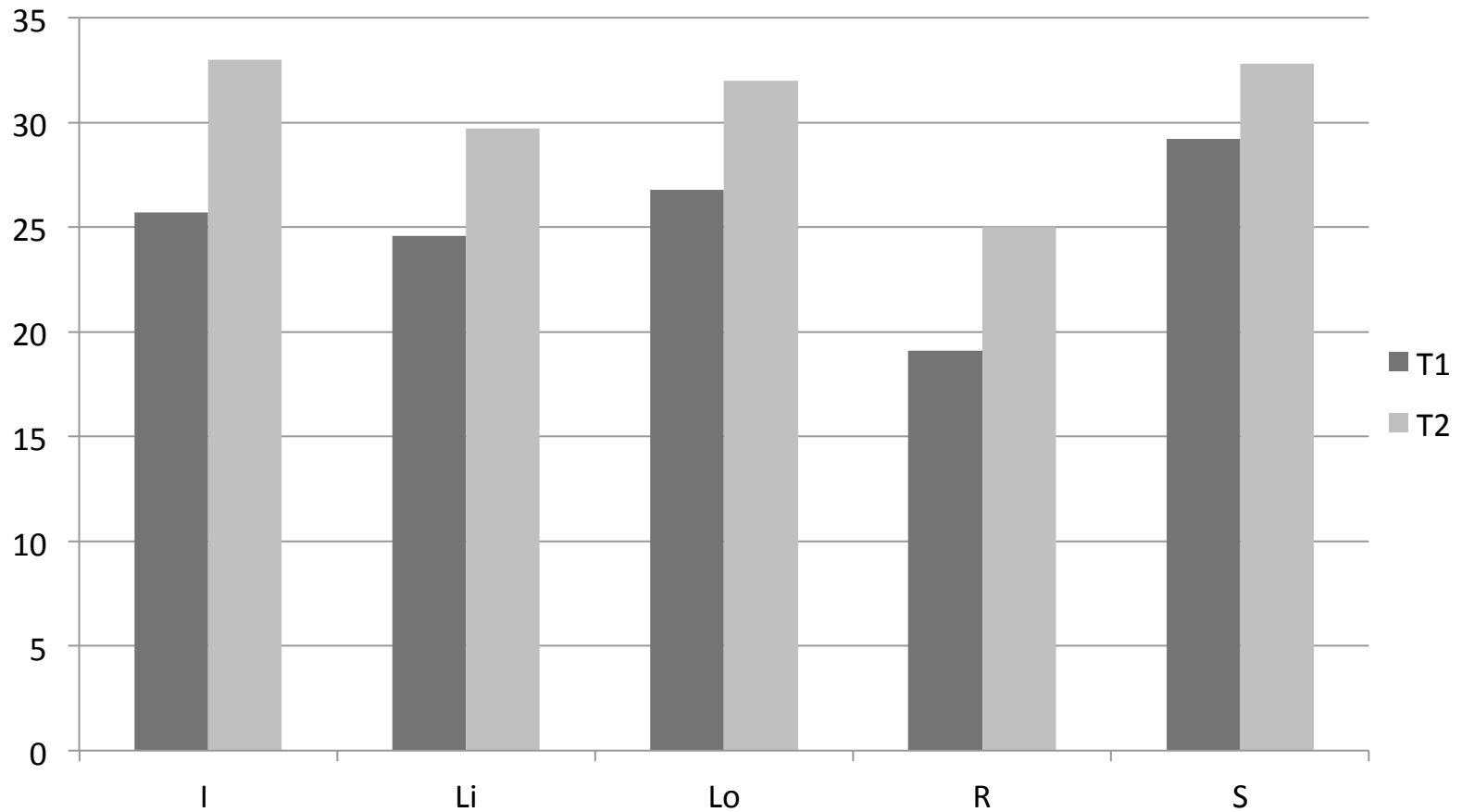
- Identification problem: determining that a sequence has been processed holistically:  
the sequence presents a processing advantage and is retrieved faster than other strings of words, but also has some kind of unitary meaning
- Example of measures used:
  - Pauses (silent and filled; over 0.2 secs)
  - Identification of fluent runs (no interruptions/pauses)
  - Application of the other identification criteria
- Research question: do advanced learners also use FS?  
If so, what is their role?

# Increase in percentage of FS use per learner between time 1 and time 2

	T1	T2	+%	P value
Iris	25.7	33	+7.3	0.13
Lily	24.56	29.66	+5.1	0.01*
Lola	26.78	31.98	+5.2	0.18
Rose	19.1	25.02	+5.92	0.06
Sally	29.18	32.74	+3.56	0.19



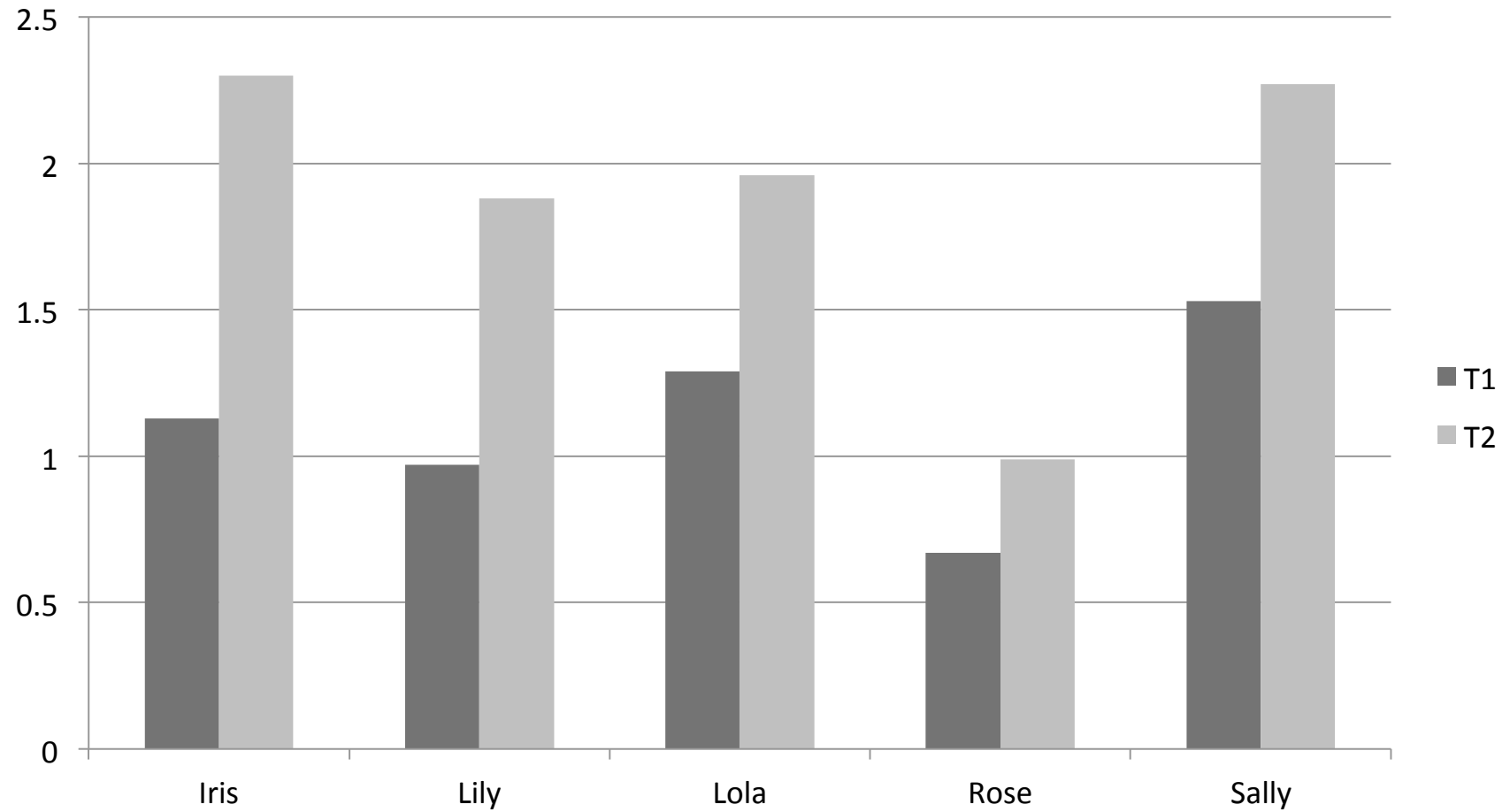
# Increase in percentage of FS per learner between time 1 and time 2



# Average no of formulaic syllables per run

	T1	T2	P value
Iris	1.13	2.3	0.0000*
Lily	0.97	1.88	0.0026*
Lola	1.29	1.96	0.0108*
Rose	0.67	0.99	0.0684
Sally	1.53	2.27	0.0060*

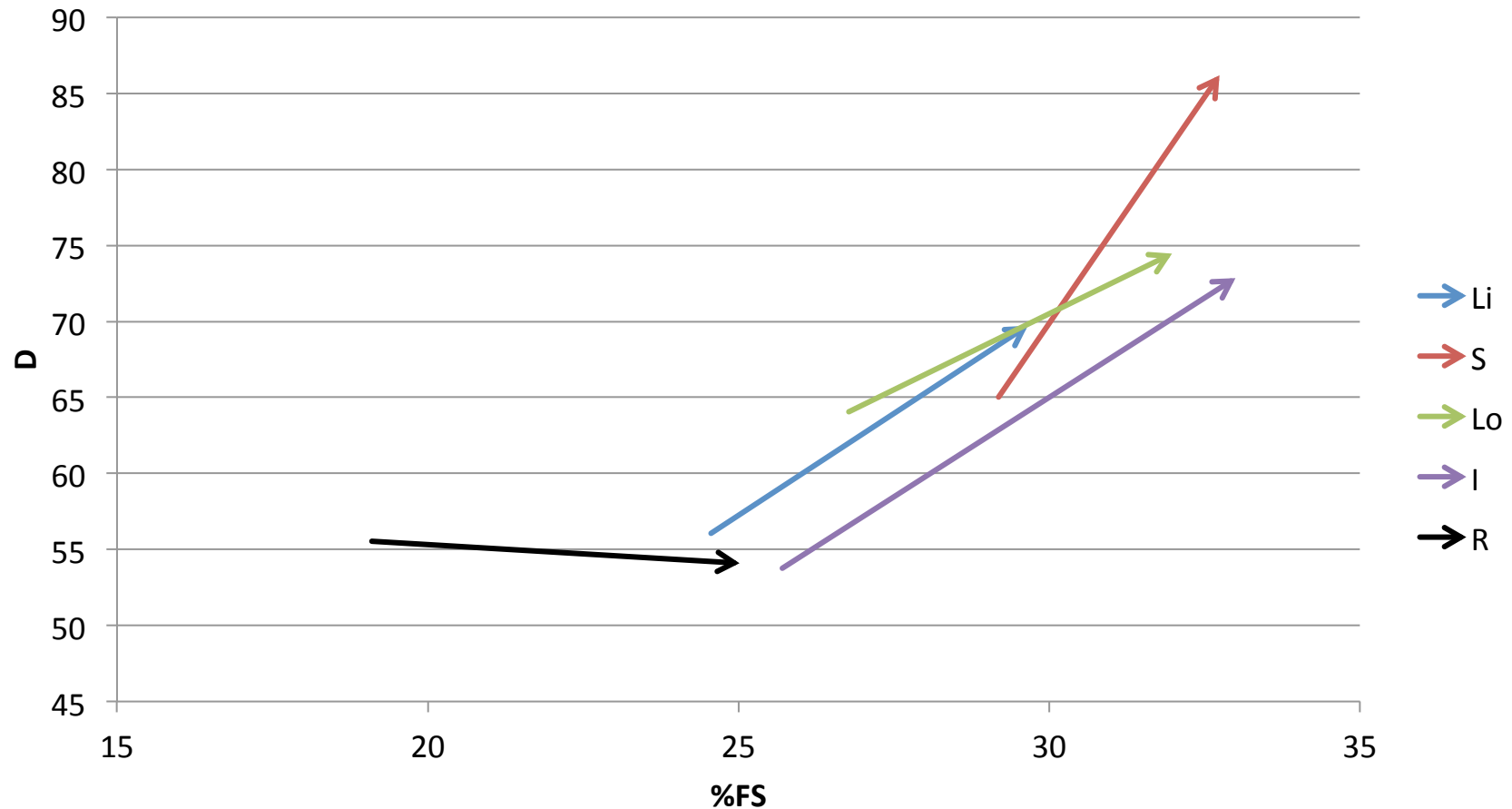
# Average no of formulaic syllables per run



# Summary development of FS during year abroad

- All learners increase their FS use significantly on at least some of the measures:
  - More FS
  - Longer FS

# Development of FS use in relation to development of lexical diversity



All learners develop the two in parallel except Rose

# Summary relationship FS use and lexical diversity

- The more a learner is formulaic at time 1, the more likely s/he is to be lexically diverse at time 2
- The automatization process involved in FS is crucial for freeing attentional resources to learn more vocabulary/grammar

# Fluency development: amount of speech

	Iris	Lily	Lola	Rose	Sally
T1	3877	3316	3631	2974	5689
T2	10020	6437	7343	4191	7028
Total	13896	9753	10974	7165	12717

No of tokens used per learner at time 1 and time 2 (same tasks):  
Learners speak considerably more at time 2

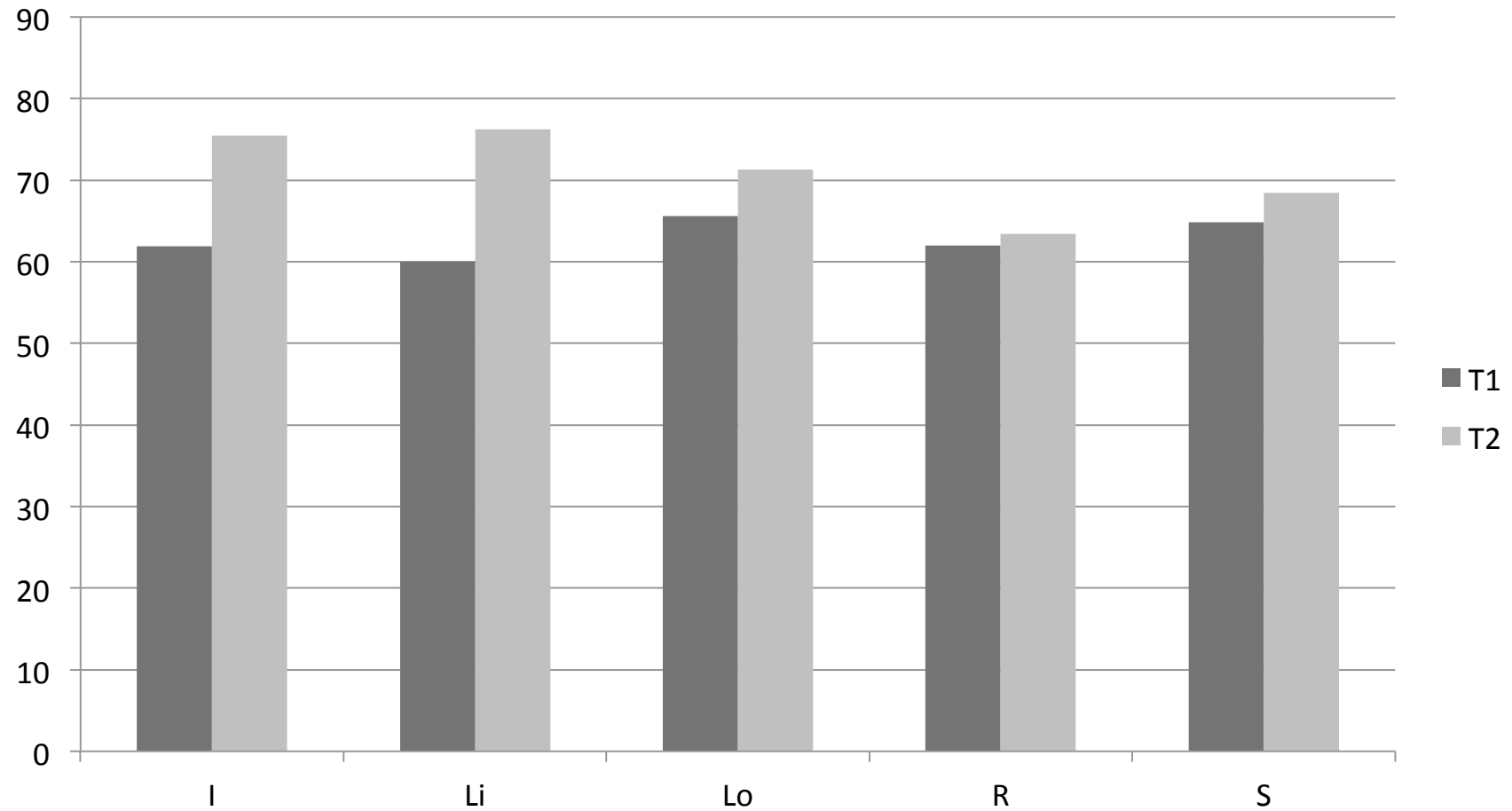
# Fluency measures: development over time

Measure	Time (df1,4)	Time x Subjects (df4,16)
PTR	26.29, p= 0.007*	10.80, p< 0.001*
MLR	250.60, p< 0.001*	9.84, p< 0.001*
SR	384.00, p< 0.001*	13.70, p< 0.001*
AR	188.52, p< 0.001*	14.25, p< 0.001*

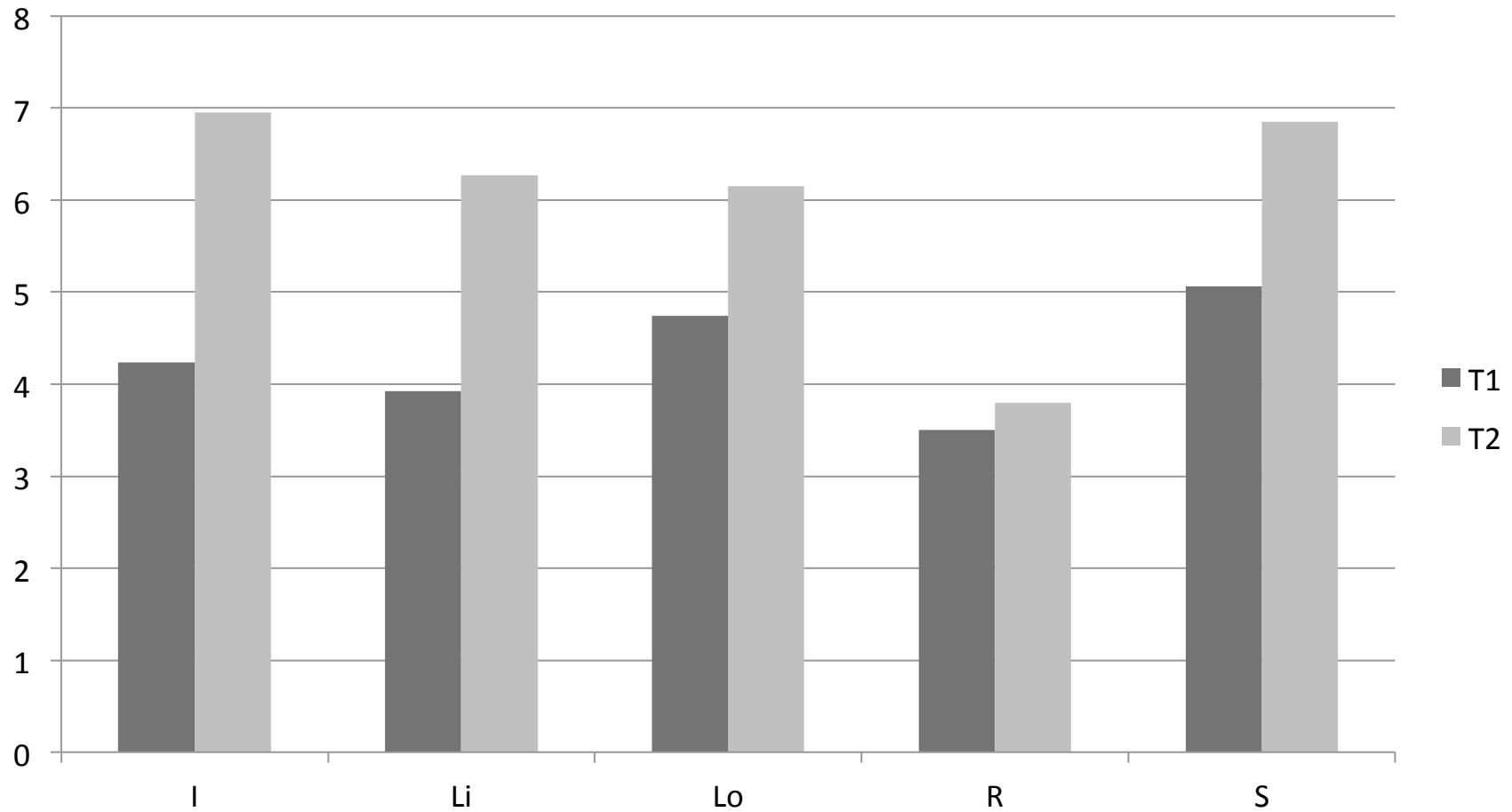
PTR: phonation/time ratio; MLR: Mean length of runs (syllables); SR: speaking rate; AR: articulation rate



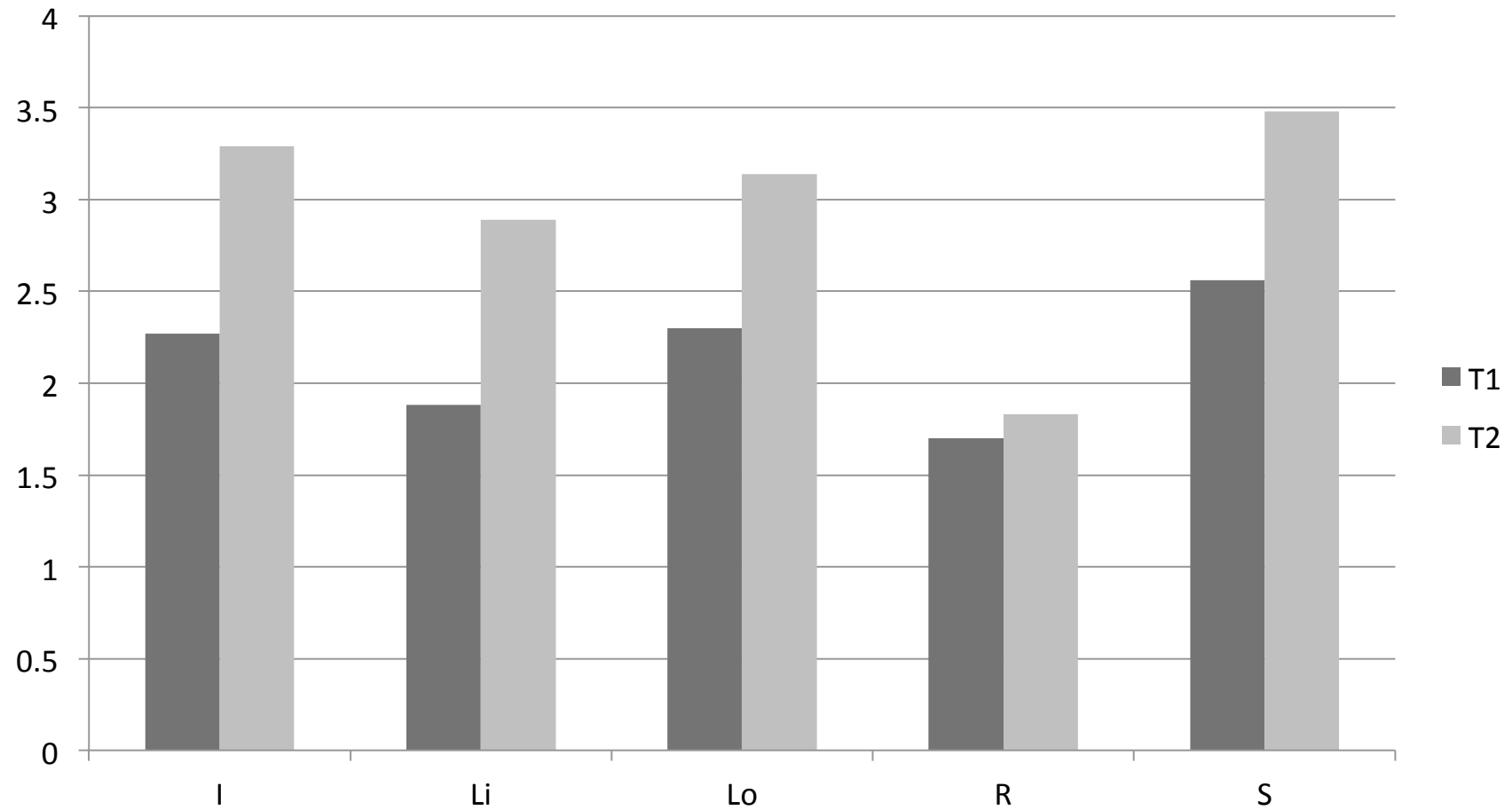
# Increase in PTR per learner between time 1 and time 2



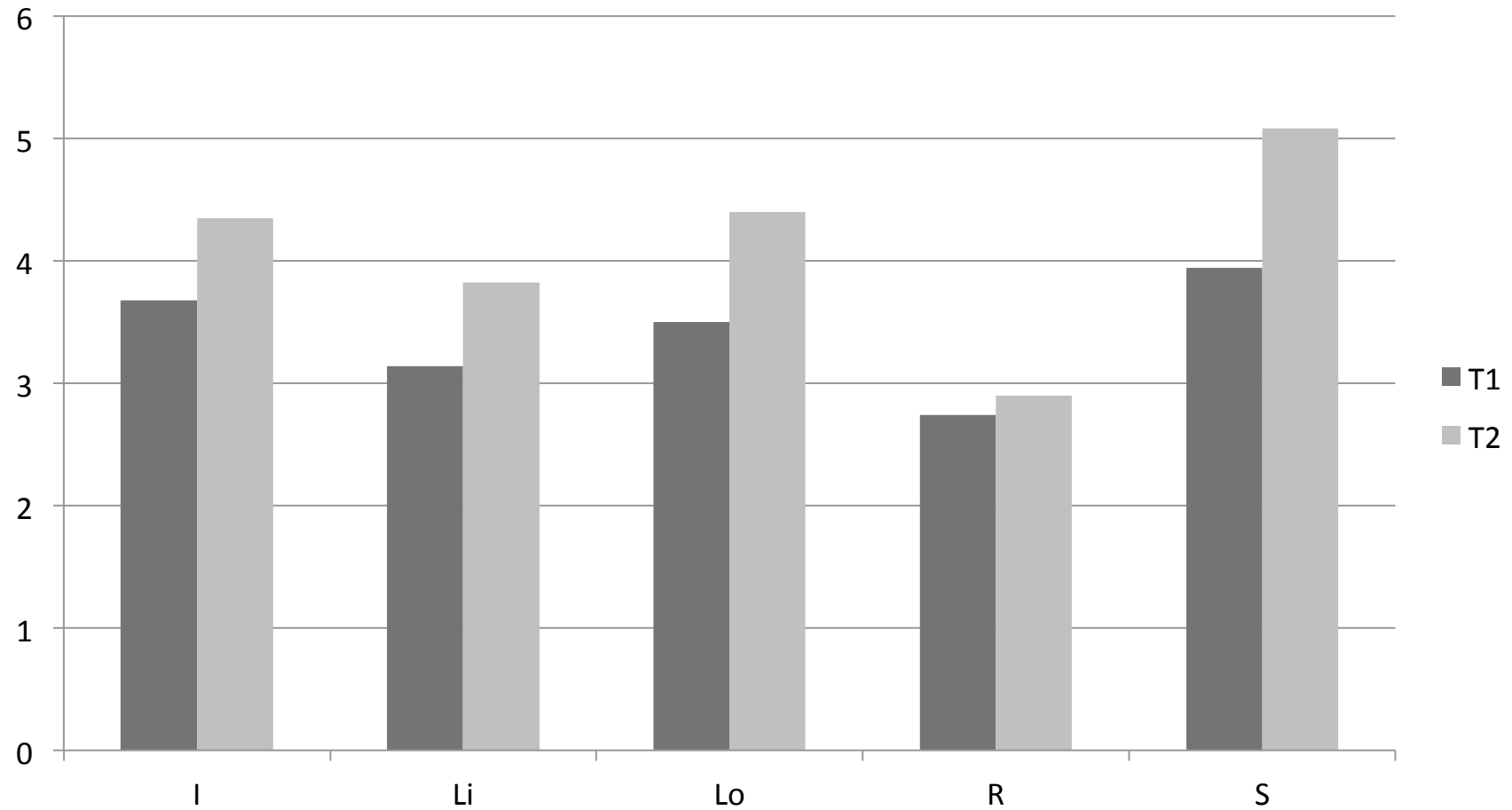
# Increase in MLR per learner between time 1 and time 2



# Increase in SR per learner between time 1 and time 2



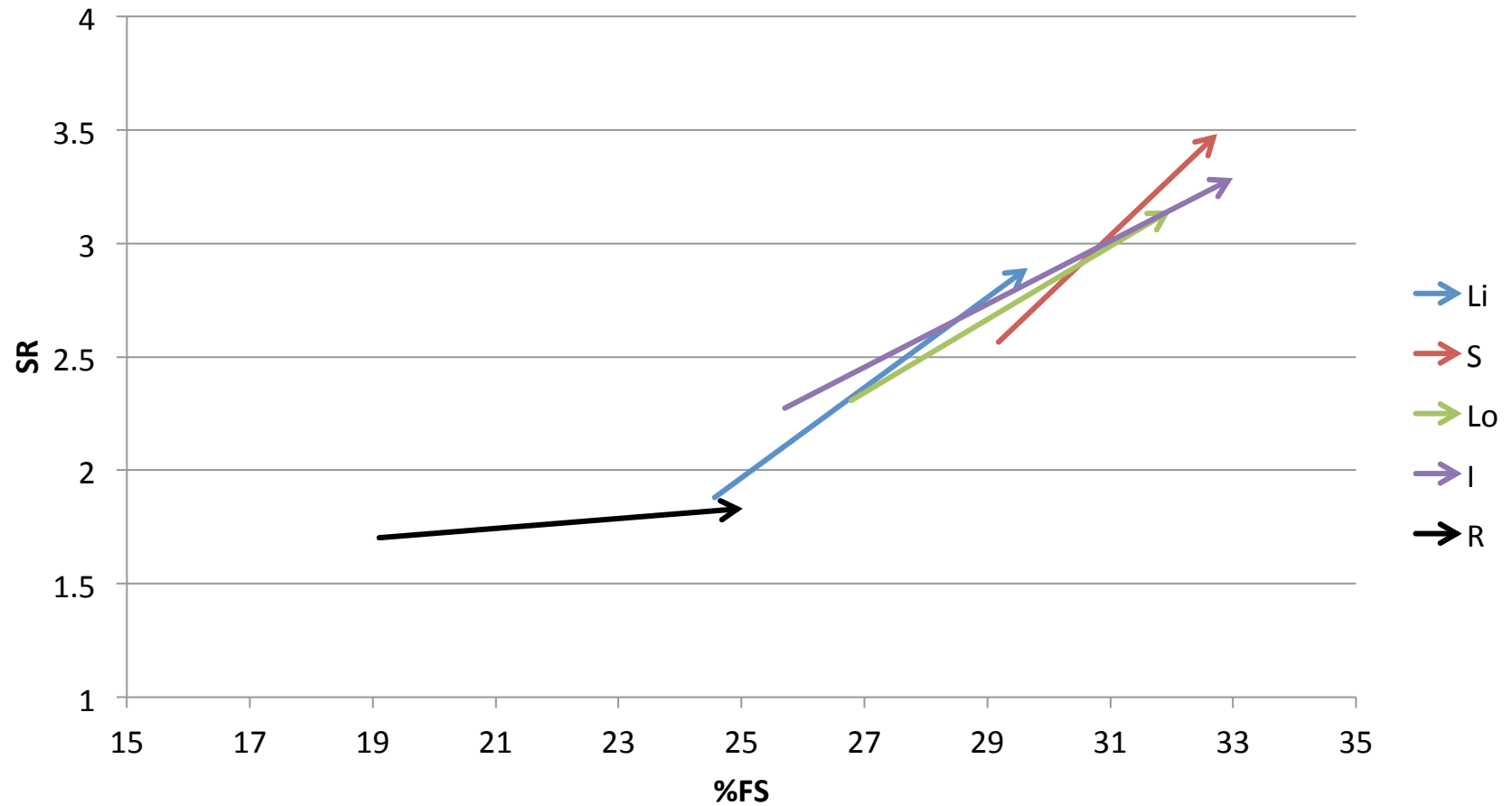
# Increase in AR per learner between time 1 and time 2



# Summary: development of fluency

- Learners speak faster, with less hesitations and pauses
- Large improvement in all learners except Rose

# Relationship between fluency and FS use



Close relationship between fluency (as measured by speaking rate) and FS use

# Summary – advanced learners

- FS do not disappear when lexical diversity increases: they become more varied and sophisticated;
- FS use and lexical diversity correlate: the more learners use FS, the more diverse their lexis is, and the more they improve on both counts over time
- FS use and fluency correlate: the more learners use FS, the more fluent they are overall
- Advanced learners still rely on the FS they learned in beginner classrooms

# Conclusion: role of FS at different stages

At both ends of the continuum:

- FS are used to facilitate speech production, but in different ways
  - Early learners use them as communicative crutches and as a model for hypothesis forming/testing
  - Advanced learners use them as processing shortcuts and to buy planning time. They also use them to free up attentional resources to pay attention to novel aspects of the language they are learning



# Pedagogical implications

- FS are useful for beginners to give them entry into communication
  - Learners should be encouraged to ‘play’ with them (i.e. modify reference; analyse their constituents...), for them to feed into their grammar building
- At more advanced levels, they facilitate processing
  - Having a stock of formulae is useful; but building this stock follows from automatisisation processes rather than explicit learning/teaching

# Some references

- Cordier, C. (2013). *The presence, nature and role of formulaic sequences in English advanced learners of French: a longitudinal study*. (PhD thesis), Newcastle University, UK.
- Ellis, N. (2012). Formulaic language and second language acquisition: Zipf and the Phrasal Teddy Bear. *Annual Review of Applied Linguistics*, 32, 17-44.
- Myles, F. (2004). From Data to Theory: the Over-Representation of Linguistic Knowledge in SLA. *Transactions of the Philological Society*, 102(2), 139-168.
- Myles, F., & Cordier, C. (in press). Formulaic sequence (FS) cannot be an umbrella term in SLA: focusing on psycholinguistic FS and their identification. *Studies in Second Language Acquisition*.
- Myles, F., Hooper, J., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, 48(3), 323-363.
- Weinert, R. (1995). The role of formulaic language in second language acquisition: a review. *Applied Linguistics*, 16, 180-205.
- Wood, D. (2010). *Formulaic language and second language speech fluency: background, evidence and classroom applications*. London: Continuum. .
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wray, A. (2008). *Formulaic language: pushing the boundaries*. Oxford: Oxford University Press.