



UNIVERSITÄT
LEIPZIG

Feuer engine, wrinkelig, and sheepe

Individual differences and analogy effects
in child bilingual morphological code-mixing

Antje Quick, Stefan Hartmann & Nikolas Koch

Funded by



Deutsche
Forschungsgemeinschaft
German Research Foundation



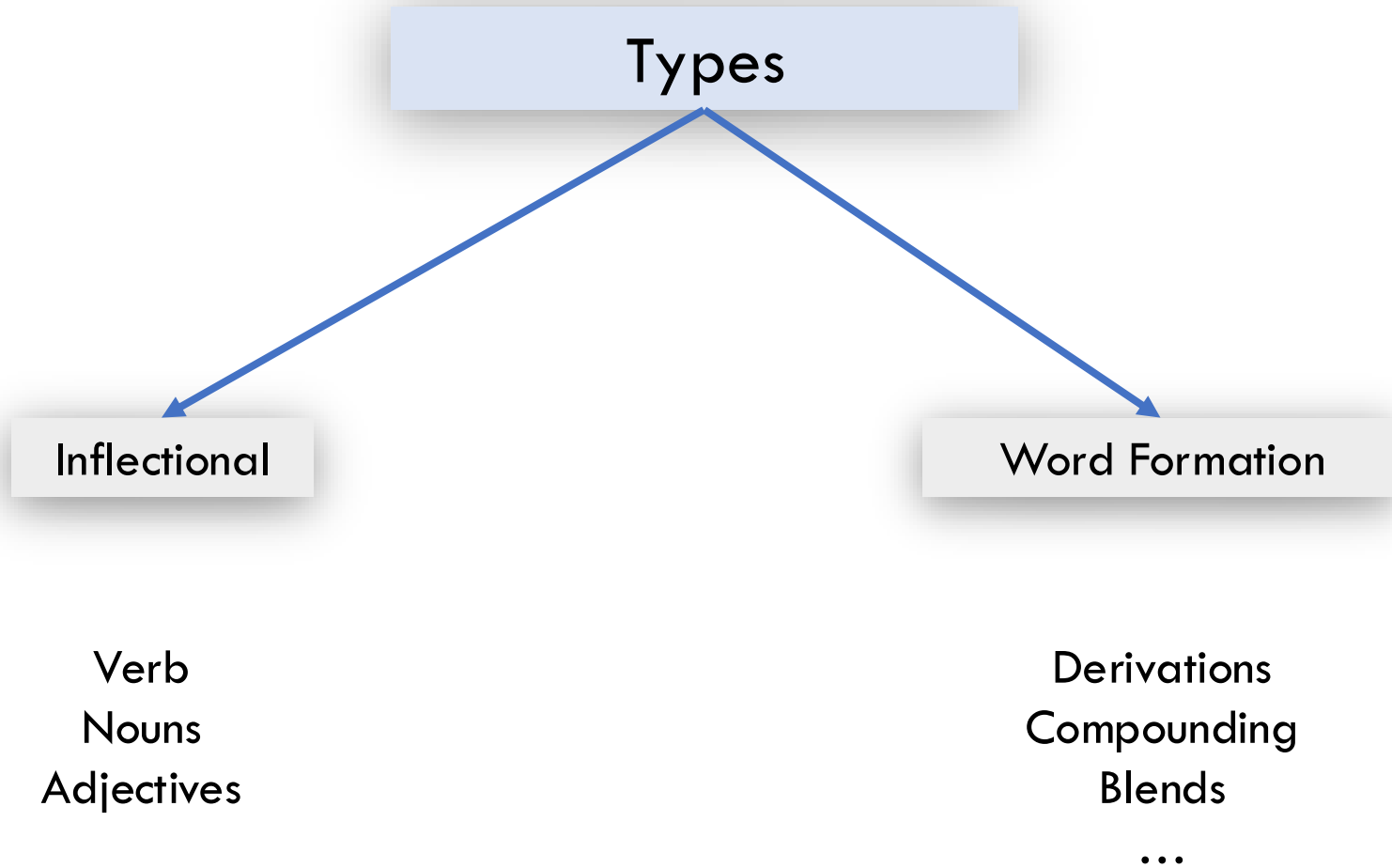
Code-mixing

Most salient contact phenomenon

- I möchte mama *“I want mama”* (Lily 2;04)
- no, a wildschwein *“no, a wild boar”* (Fion 3;05)
- du bist naughty *“you are naughty”* (Silvie 3;00)
- **cleanen** the bath *“to clean the bath”* (Fion 2;05)



Morphological Code-mixing



milken the cow

→ milk + -en (German infinitive)

ich **sleepe** bei dir

→ sleep + e (2nd PS Singular marker in German)

I'm **kaufing** this

→ buying (kauf + English present participle)

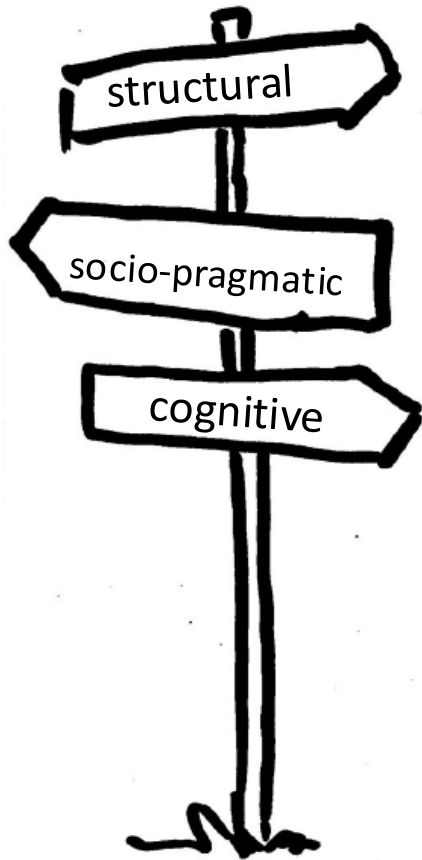
das ist wie **micen**

→ mice + -en (German plural marker)

Word Formation

polizei car	→ police car (Compound)
police frau	→ policewoman (Compound)
wrinkelig	→ wrinkle + -ig (German adjective marker)
moond	→ moon + mond (blend)
squirrel lied	→ squirrel song (Compound)
outseite	→ outside/Außenseite (Compound)
scheig	→ show + zeig (blend)

Approaches to code-mixing



and many more...

Structural Approaches → prioritize grammatical rules

Socio-Pragmatic Approaches → focus on Social motivations and conversational functions of code-mixing

Psycholinguistic Approaches → focus on activation and access of the two languages in the bilingual mind

Cognitive/Usage-based Approaches → language structure as emerging from language use, role of cognitive processes, e.g. analogy

Usage-based approach



- language learning and use as part of broader cognition
- language is learnt through actual language use and exposure to input
- language emerges and is shaped by general cognitive processes, e.g. memory, pattern recognition, attention, analogical reasoning, ...

Analogy

- key role in explaining how speakers create, extend, and modify linguistic patterns based on experience

[X_ed]

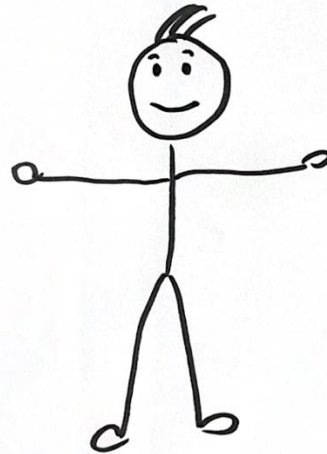
walk – walked
play – played

cheer – cheered
go – *goed

moon – mond

moond

/mu:n/ - /mo:n t/



Research Questions

Which types of word-internal code-mixing can be found in the two children's data?

Which individual differences can we find between the two children?

Does the children's input situation predict the types of word-internal mixes that we find?

Can we find analogy effects (e.g. phonological proximity to cognates) in word-internal mixes?

Data - Bilingual Corpora

DE GB

Silvie

2;4 - 3;10

$N_{CHI} = 65.473$

$N_{Input} = 140.387$

DE GB

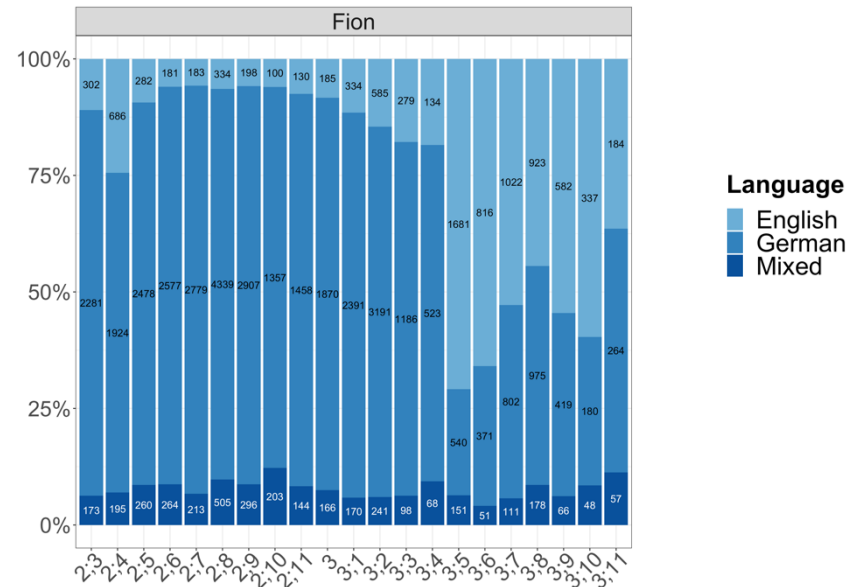
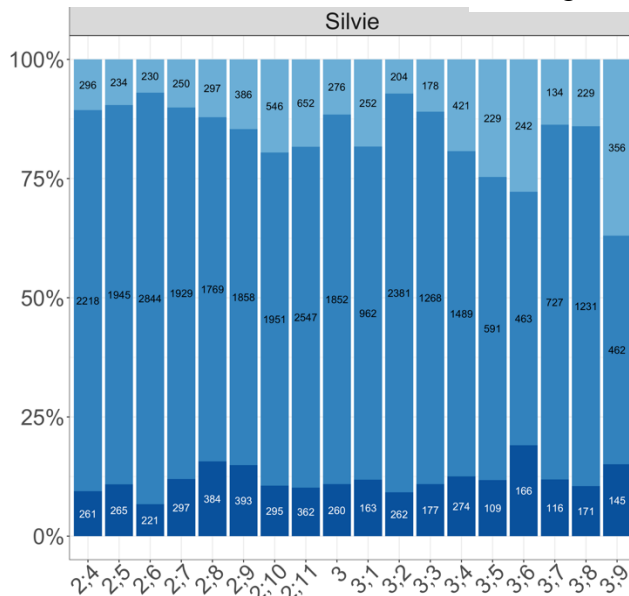
Fion

2;3 - 3;11

$N_{CHI} = 47.928$

$N_{input} = 180.292$

Language Proportions



Data - Bilingual Corpora

DE GB

Silvie

2;4 - 3;10

$N_{\text{CHI}} = 65.473$

$N_{\text{Input}} = 140.387$

DE GB

Fion

2;3 - 3;11

$N_{\text{CHI}} = 47.928$

$N_{\text{input}} = 180.292$

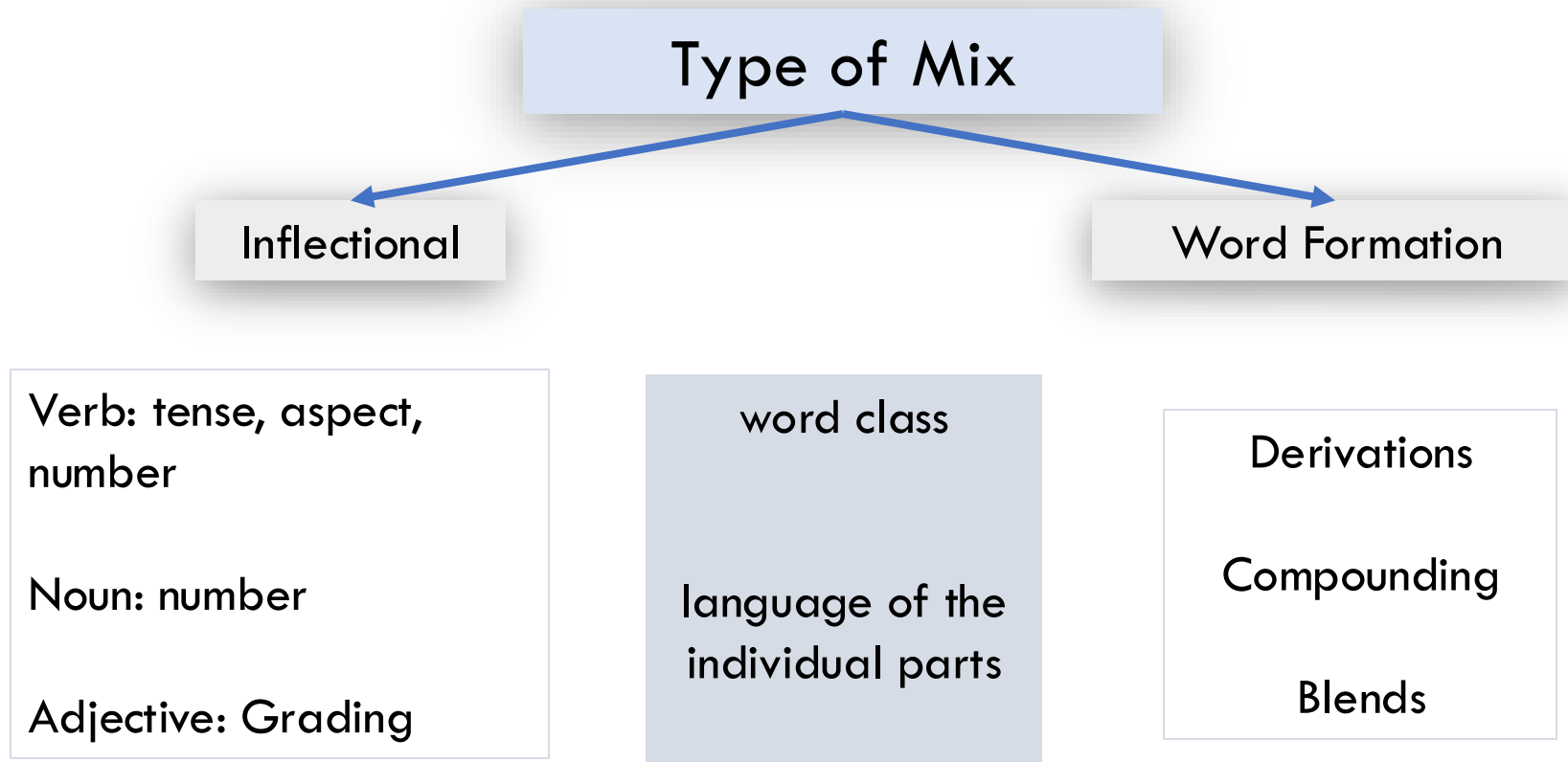
Code-mixed utterances: 4293

Word-internal mixes: 370

Code-mixed utterances: 3515

Word-internal mixes: 205

Method – coding word-internal mixes



tickeln → inflectional mix → engl. V + German infinitive marker

wrinkelig → derivational mix → engl. Adj + German derivational suffix (-ig)

fertnished (finished – fertig) → blend → German + English

Doubtful cases

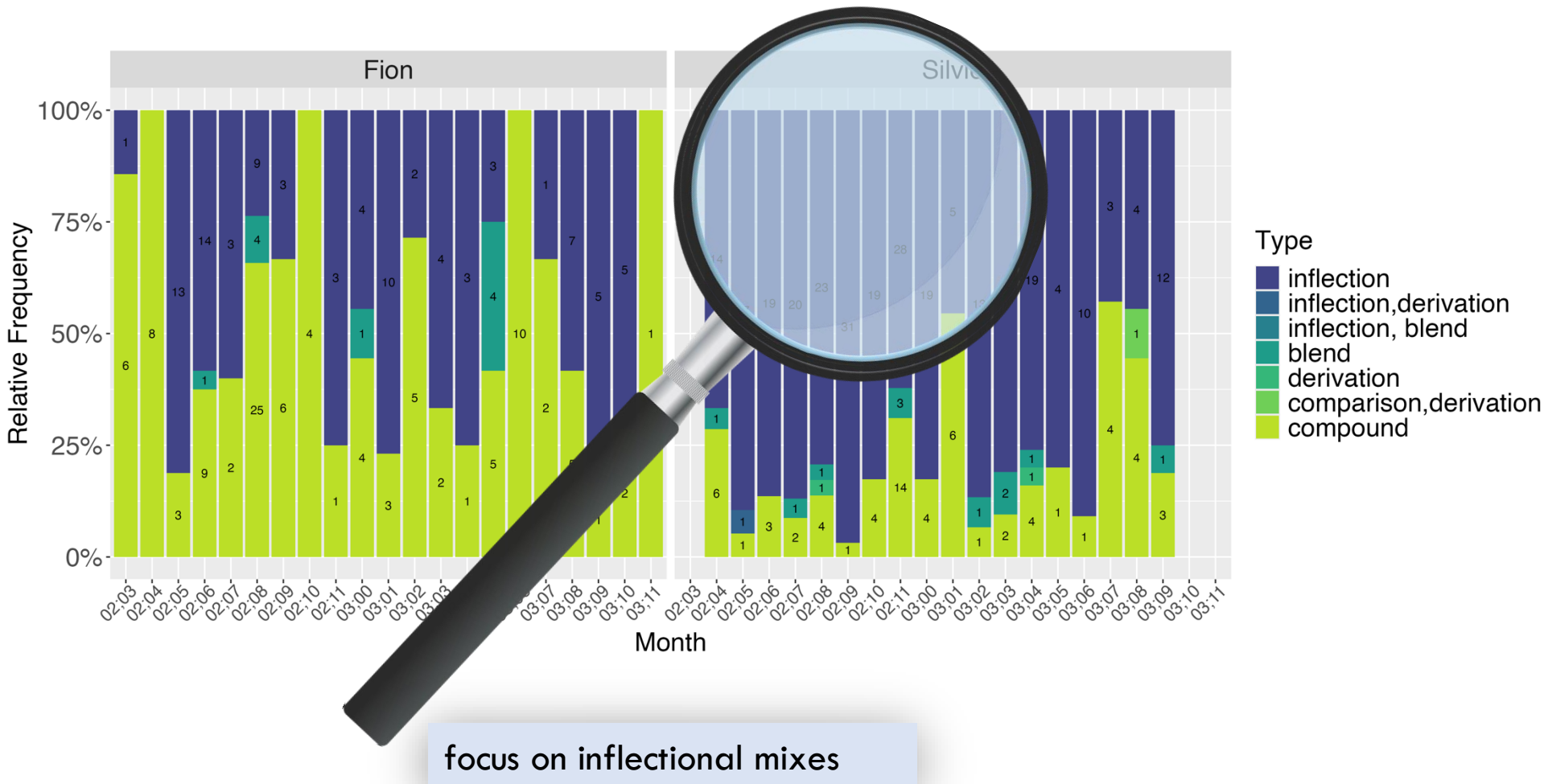
- words with a silent <e> in their orthographic representation, e.g. <squeeze> (/skwiːzə/ or /skwiːzθ/)
 - (in some cases, transcribers added an additional <e> for clarification, e.g. <bounceee>, but apparently this has not been done systematically → words ending in single silent <-e> in English standard orthography excluded)
- cross-linguistic (near-)homophones, e.g. <mami bear> ≈ <mommy bear> → omitted for the present study
- words that also exist as loan words in German, e.g. *killen* 'kill', *kicken* 'kick' → omitted as well (exception: *scraper* 'scrape', exists as a loan word as well but only as a technical term)

Frequency of word-internal mixes

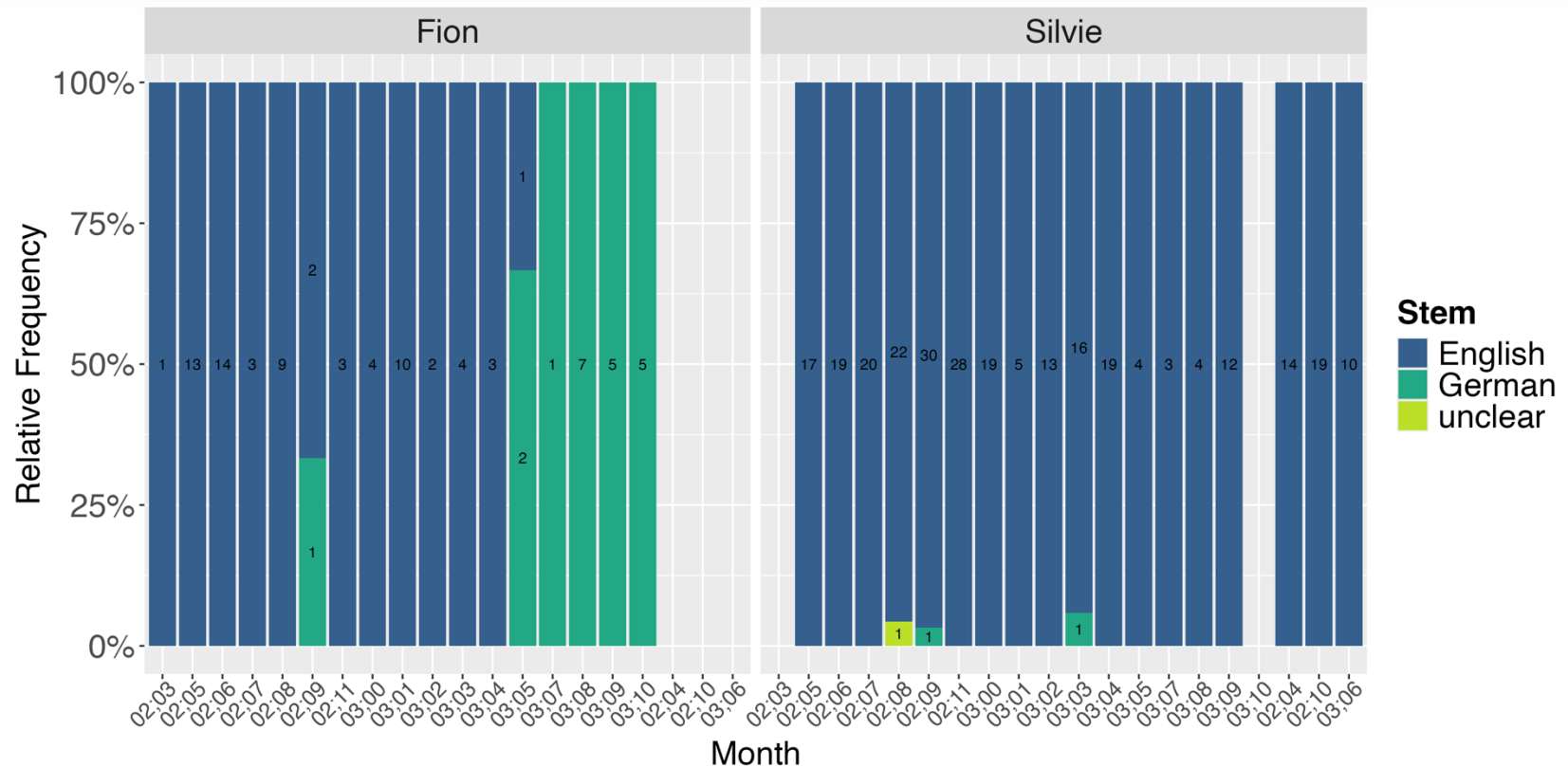
of word-internal mixes / # of utterances



Types of word-internal mixes



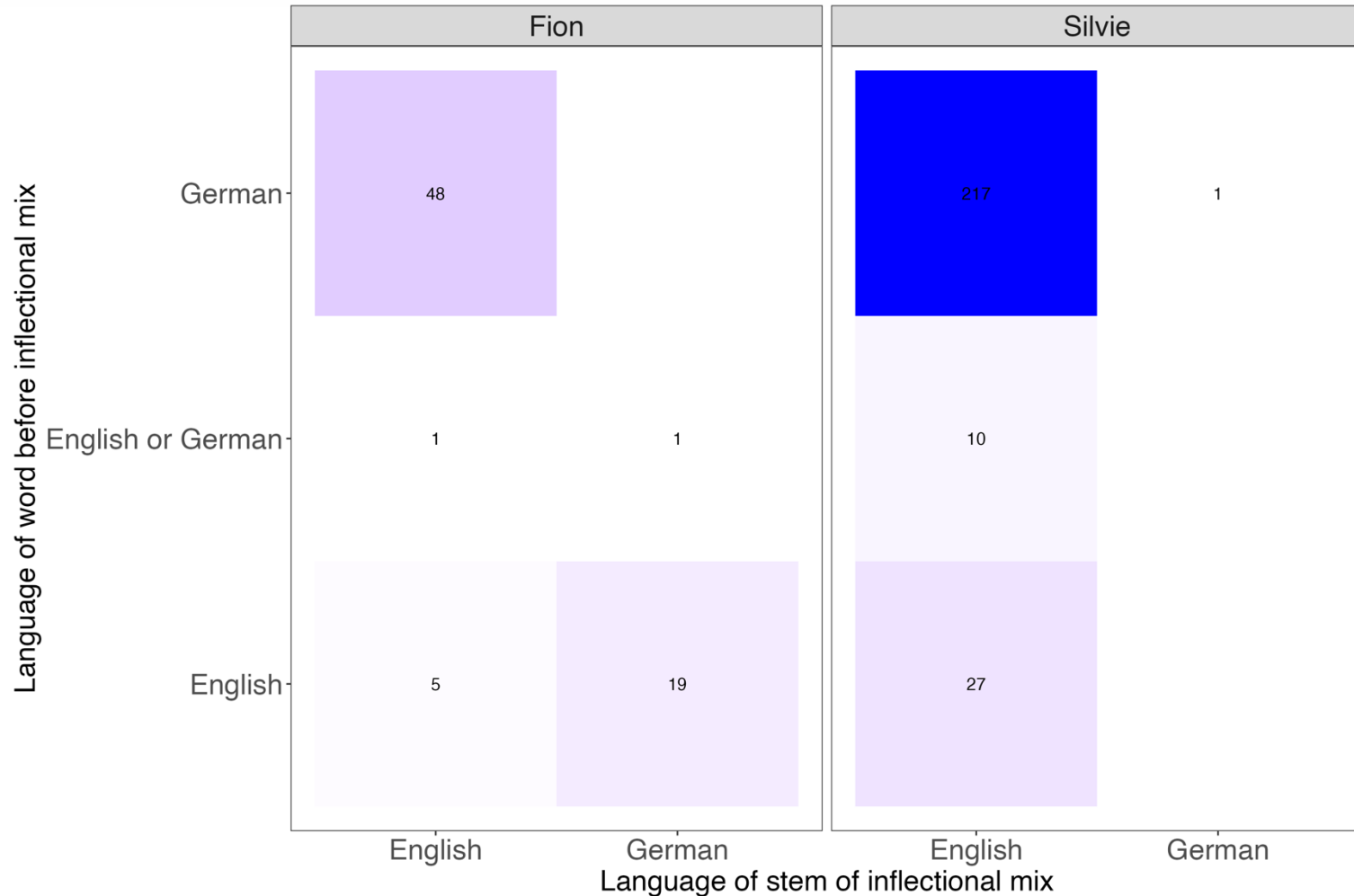
Language of stems in inflectional mixes



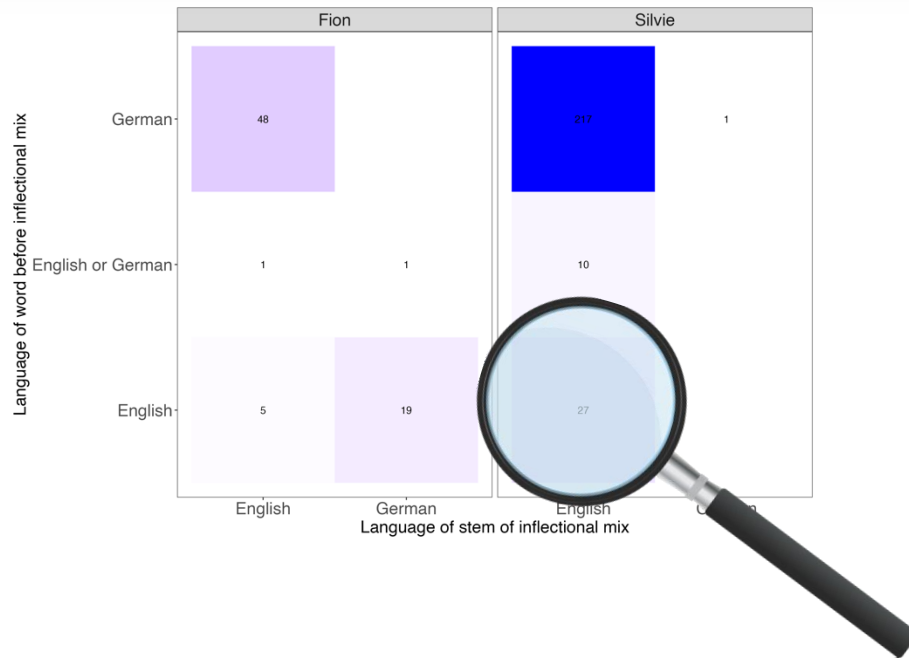
musik hearen – to hear music+ en (hören)
 der sleep immer - he always sleeps (schläft)

he had no beins – we had no legs (beine)
 we're baving a thing – we're building a thing

Inflectional mixes: Language of preceding word

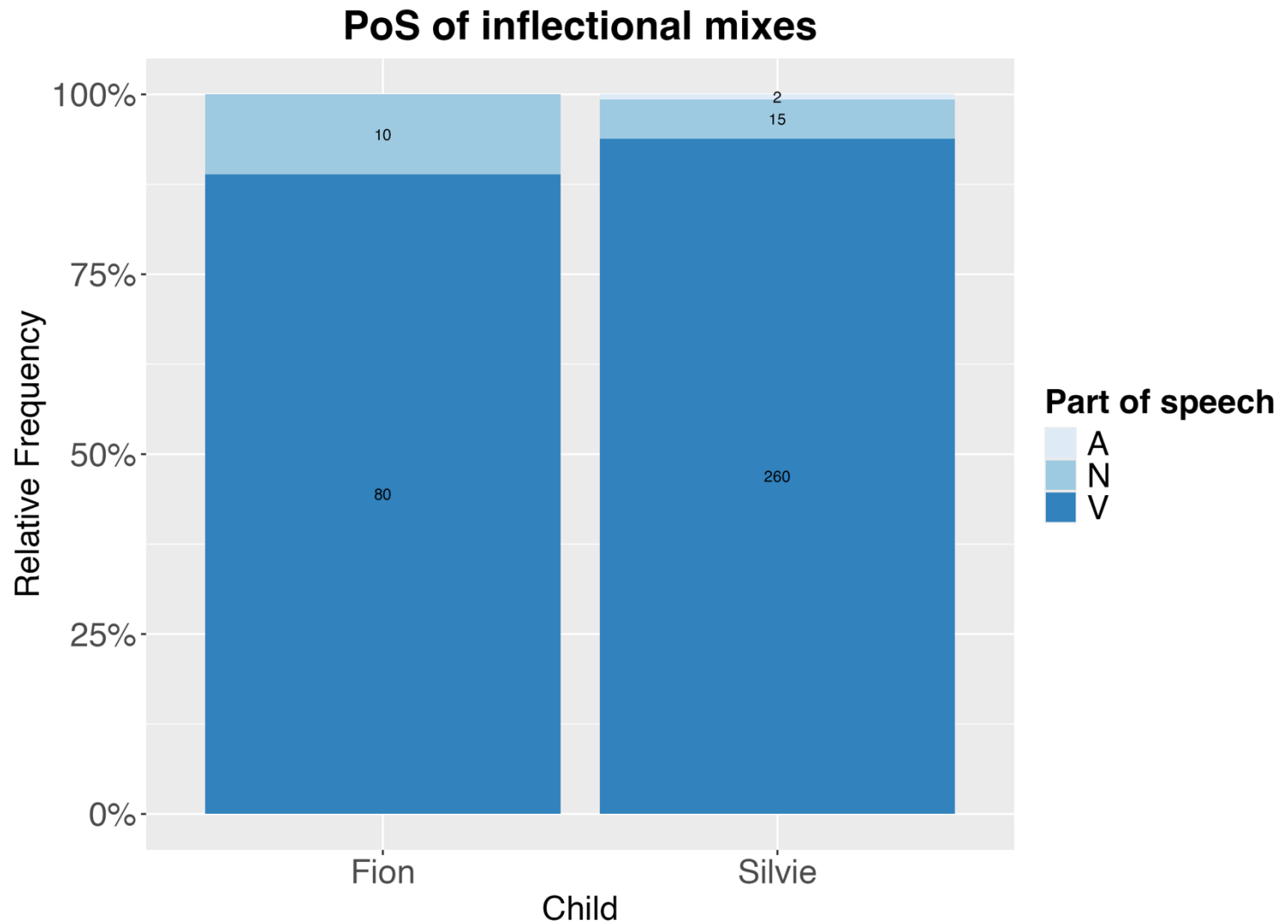


Inflectional mixes: Language of preceding word

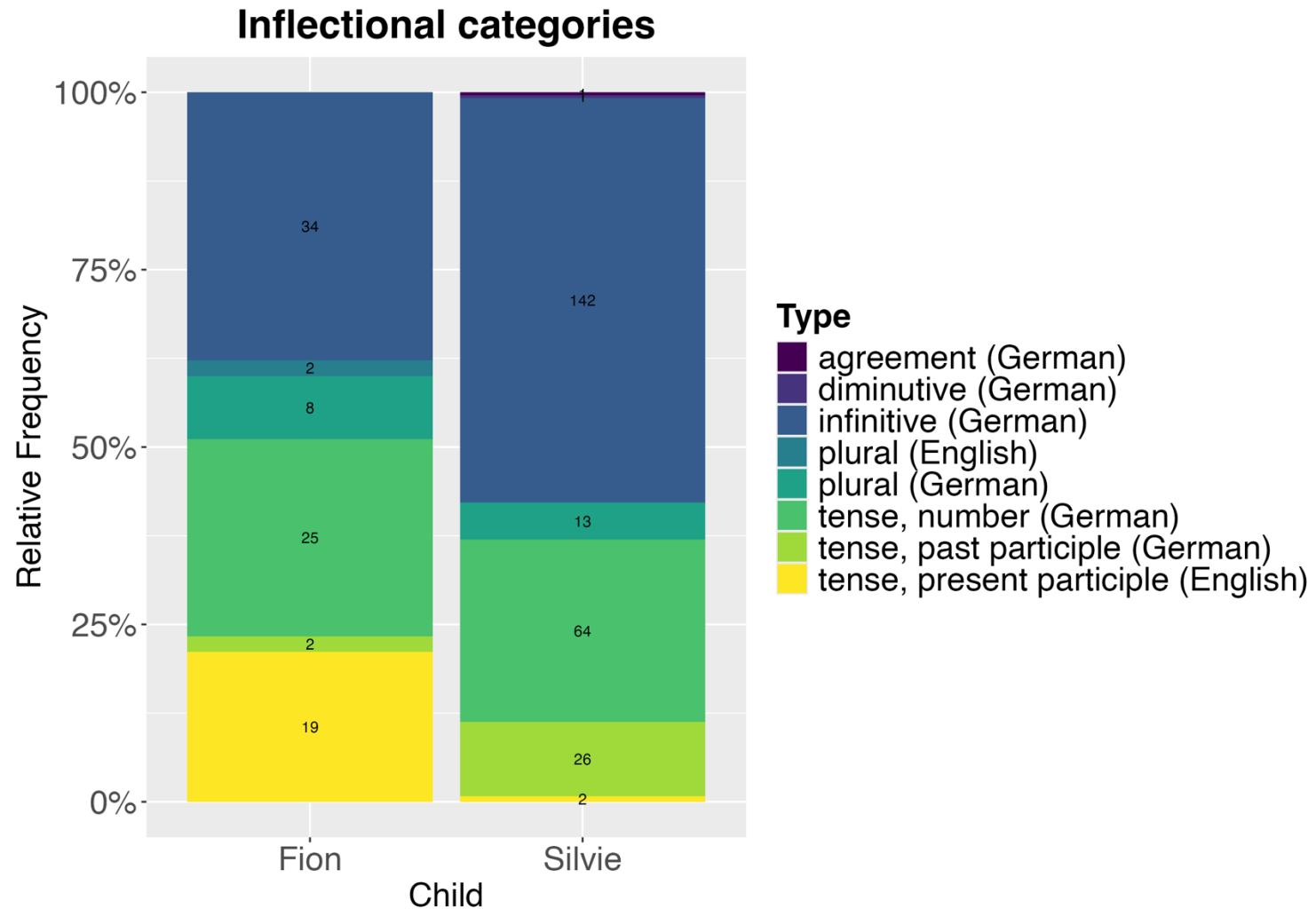


- the lions **moven** sich schon out of the way – the lions move themselves out of the way (move + en → gehen)
- aber in the mountains **liven** auch squirrels - but squirrels live in the mountains too (live + en → leben)
- chicken **cooken** wir – we cook chicken (cook + en → kochen)

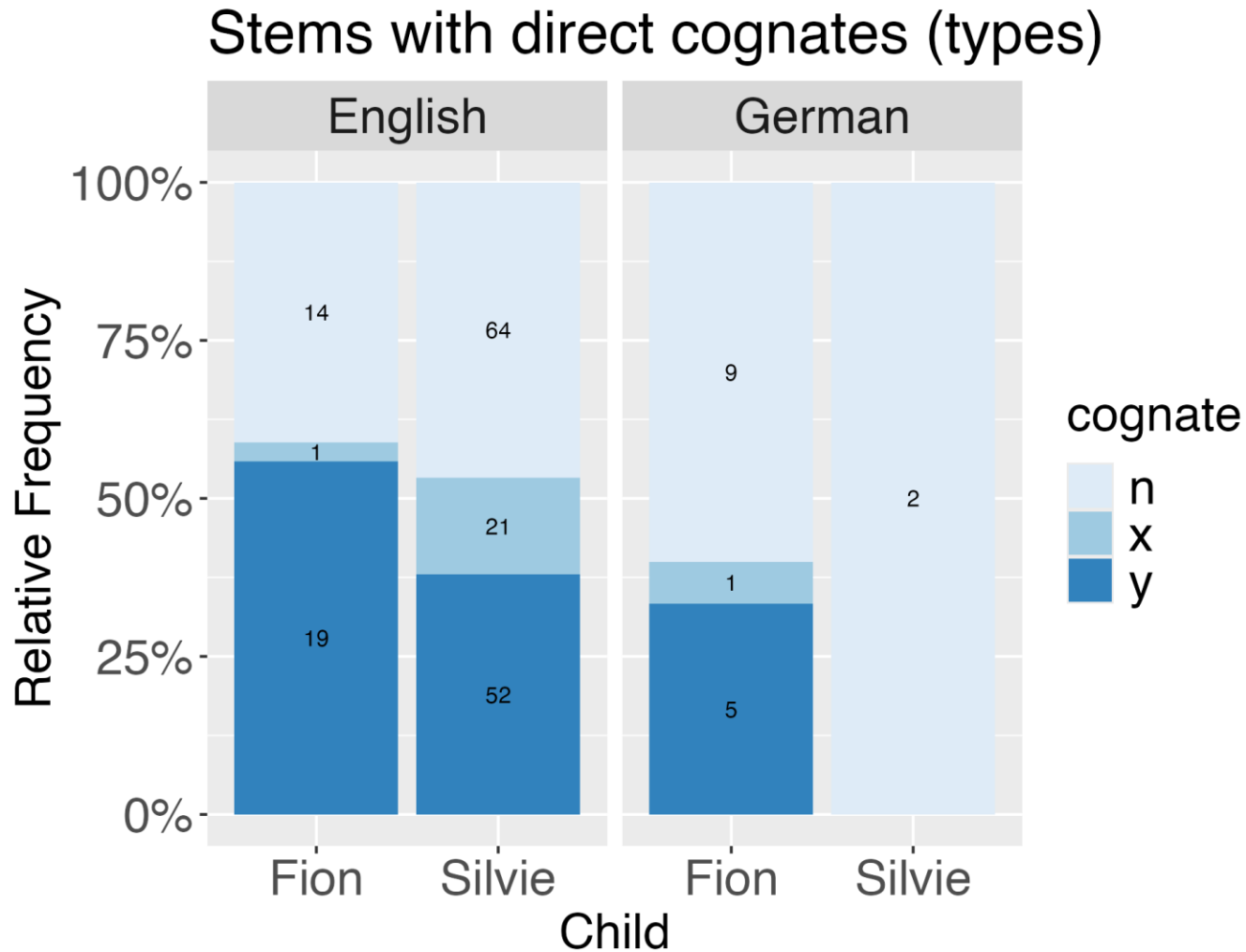
Inflectional mixes: PoS



Inflectional mixes: Categories



Inflectional mixes: Analogy to cognates?



Levenshtein Distance

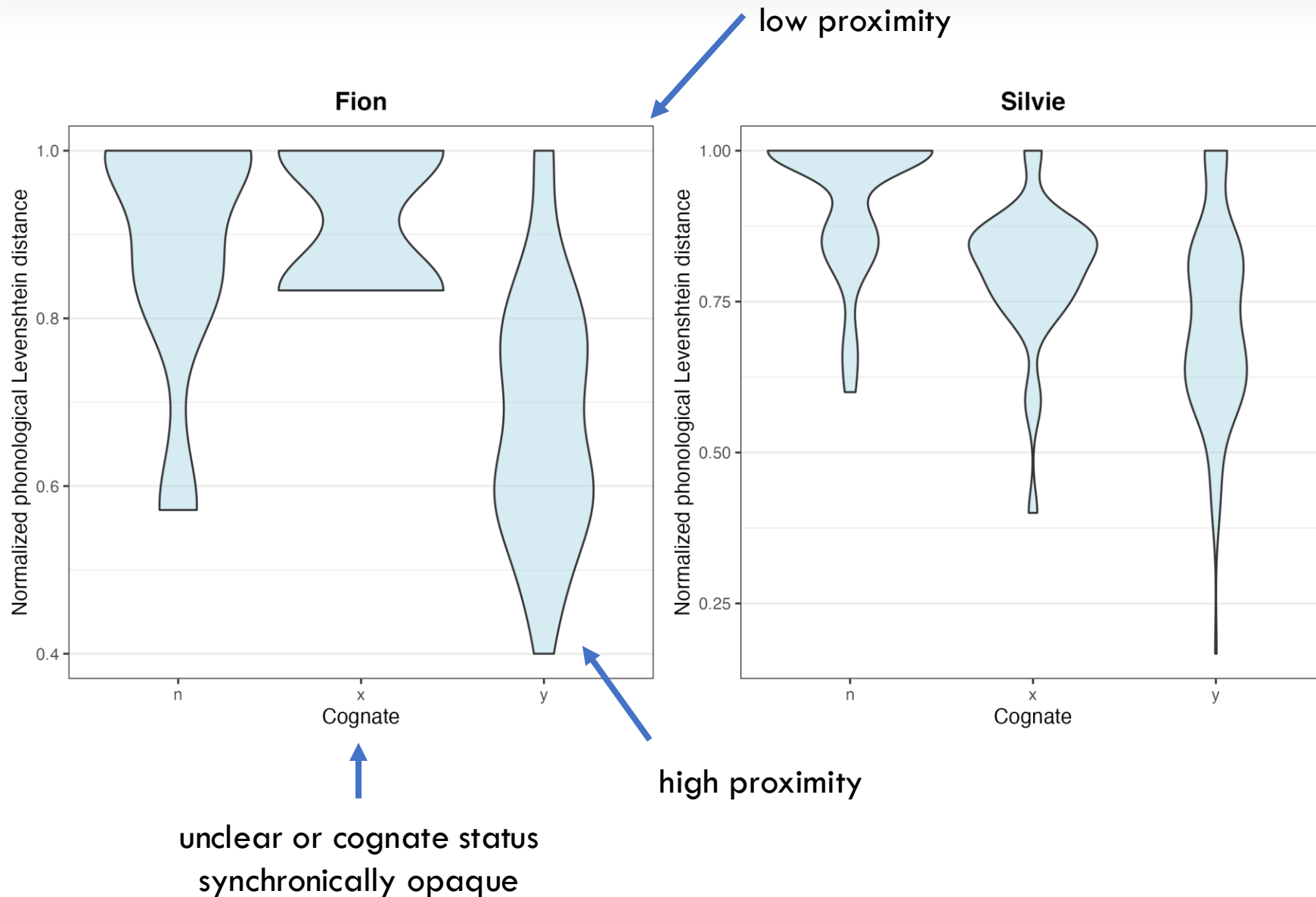
- numerical, objective way to quantify linguistic differences → how similar/different two words are
- moond → /mo:ɪn t/ → /mu:ɪn/ → Levenshtein Distance = 2

Step	From	To	Operation
1	o:ɪ	u:ɪ	substitution
2	t	(nothing)	deletion

- sleepe → /sli:p/ vs /ʃlaɪfə/ → Levenshtein Distance = 4
- cleanen → /kli:n/ vs /ʁaɪnɪgən/ → Levenshtein Distance = 7
- data automatically transcribed to ARPABET via G2P (BAS Web Services, LMU Munich)
- normalized Levenshtein distance: divided by length of longer item → $0 \leq \text{score} \leq 1$

Lower scores = more similar pronunciations

Results



Lemmas in inflectional mixes

	Fion		Silvie	
Rank	Lemma	n	Lemma	n
1	cry	7	sleep	11
2	drive	6	squash	9
3	tickle	5	eat	8
4	clean	3	bake	7
5	go	3	squeeze	7
6	kaufen	3	sheep	6
7	pick up	3	stir	6
8	sleep	3	bounce	5
9	speak	3	burn	5
10	walk	3	help	5

Research Questions

Which types of word-internal code-mixing can be found in the two children's data?

Which individual differences can we find between the two children?

Does the children's input situation predict the types of word-internal mixes that we find?

Can we find analogy effects (e.g. phonological proximity to cognates) in word-internal mixes?

Discussion

- individual differences in the types of morphological mixes → Silvie more inflectional, Fion more compounds
- especially mixed stems with direct cognates seem to be influenced by phonological proximity
- especially inflectional mixes can be accounted for by lexically fixed patterns and emerging frame-and-slot patterns → *spreaden* [X_en] German infinitive marker
- generating/recycling (creative) utterances from ‘bits and pieces’ of already acquired constructions
- children generate their code-mixed utterances directly from highly frequent strings

Спасибо Gracias شكر Obrigado Спасибо Dank U
Grazie Euxαριστώ Danke
Merci Thank You Ngiyabonga Dank U
Diolch Thank You
Dank U
Dziękuję Dziękuję
Danke Grazie
Merci Dank U
Terima Kasih
Tack Tack
Euxαριστώ