

UiT

**THE ARCTIC
UNIVERSITY
OF NORWAY**

Может ли языковой контакт способствовать морфологическому упрощению? Результаты экспериментального моделирования

Александр Бердичевский (по материалам совместного исследования с Артуром Семенюком, Калифорнийский университет в Сан-Диего)

Открытый семинар ИЛ РГГУ, 28.05.2015



План

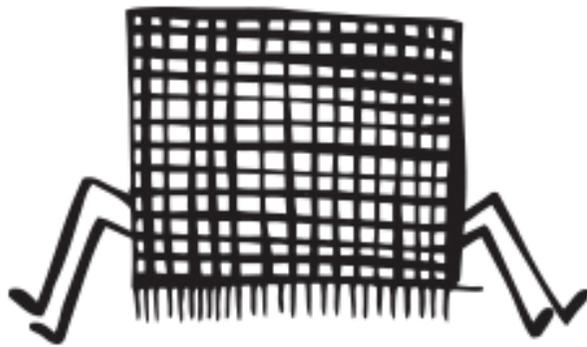
- Дизайн эксперимента (глазами участника)
- Теория, смысл и предсказание (кратко!)
- Дизайн эксперимента (глазами лингвиста)
- Результаты и их анализ
- Обсуждение

ДИЗАЙН ГЛАЗАМИ УЧАСТНИКА

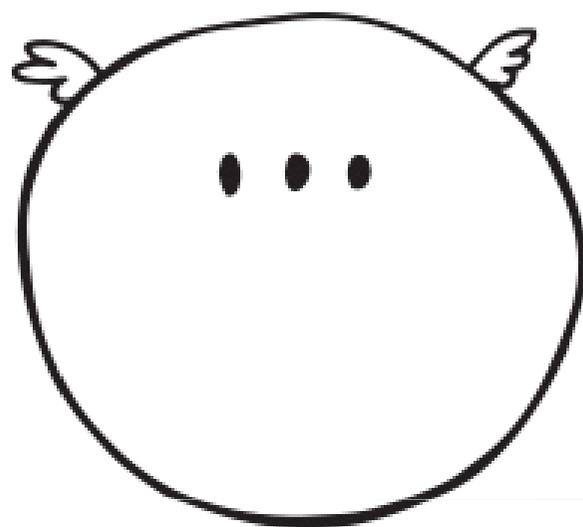
Вы участвуете в космической экспедиции, налаживающей контакт с обитателями далекой планеты Эпсилон. Эпсилонцы приветливы и охотно показывают свою планету. Сегодня эпсилонец по имени Сеуссе хочет научить вас своему языку.

Он будет показывать вам фотографии эпсилонских животных и описывать их по-эпсилонски. Время от времени Сеуссе будет проверять, как вы усваиваете язык.

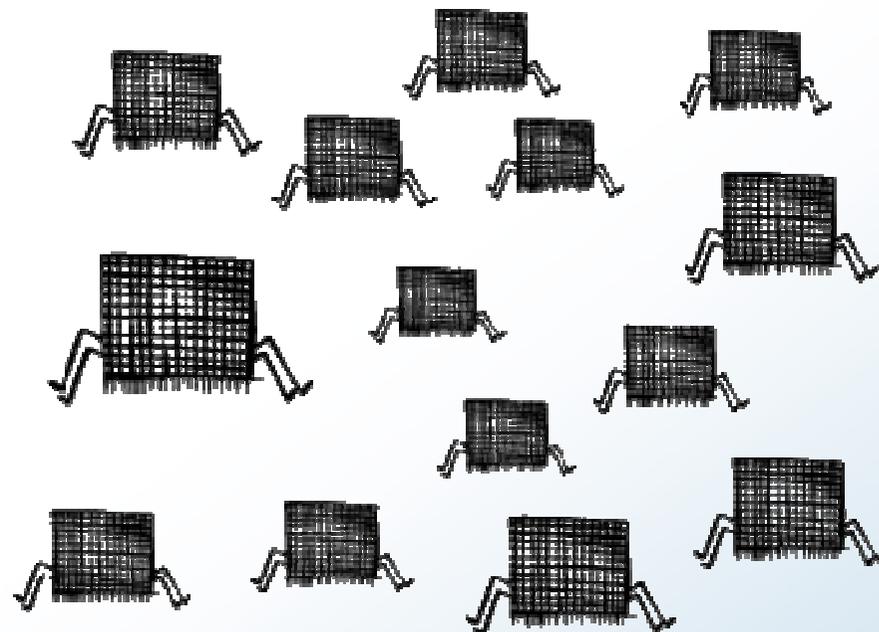
Не расстраивайтесь, если вам покажется, что у вас не очень получается: ваш друг будет изо всех сил стараться вас понять.



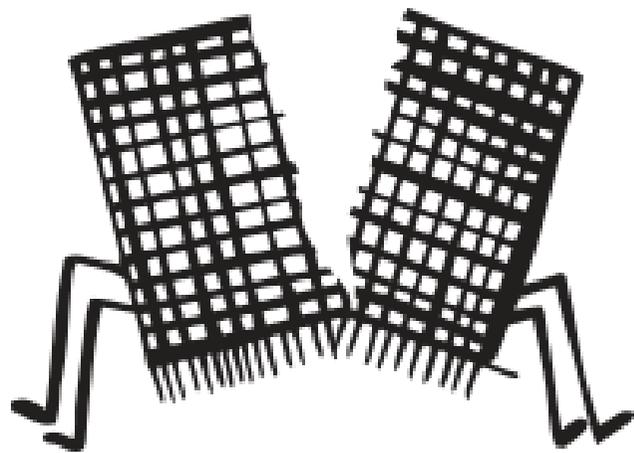
fuv



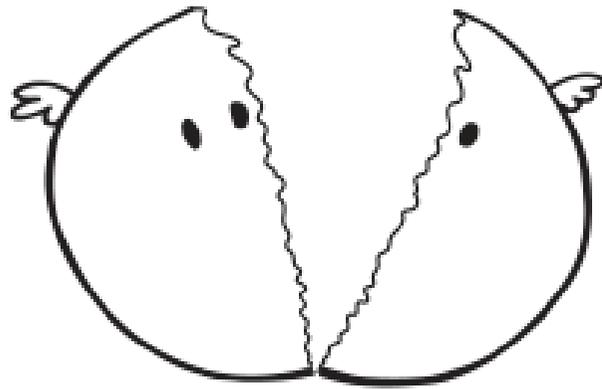
seg



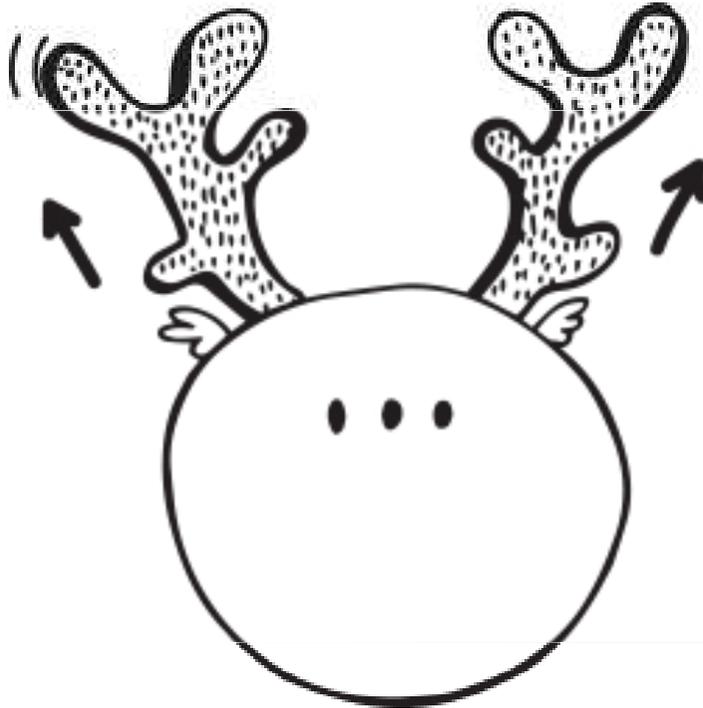
fuvl



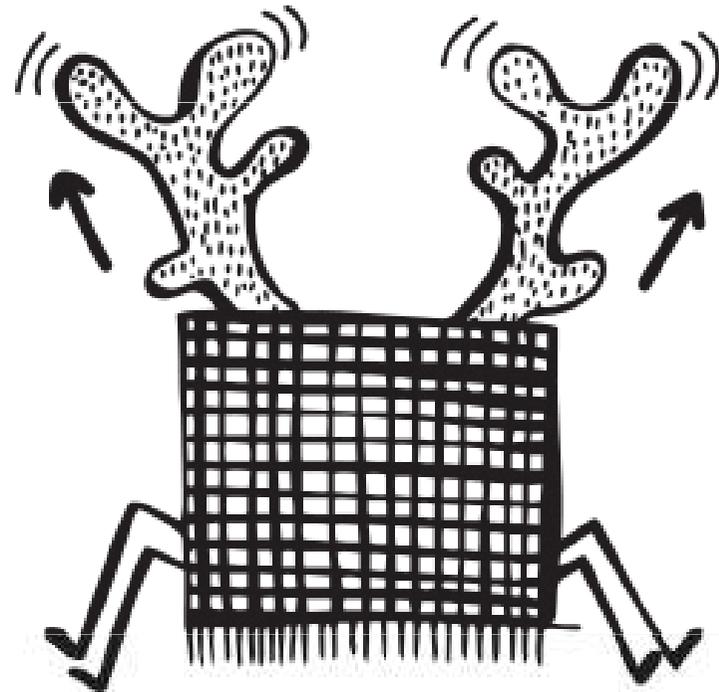
fuv mi



seg mo



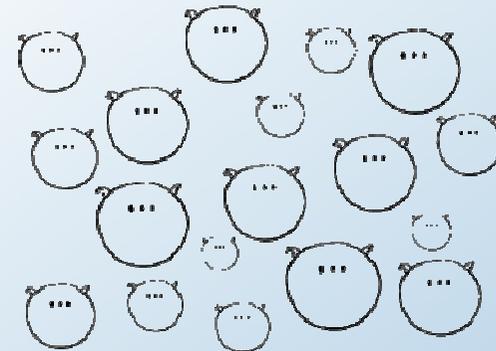
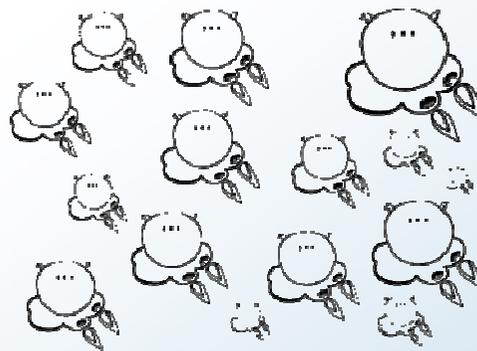
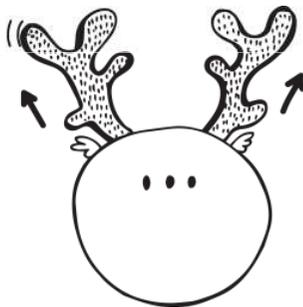
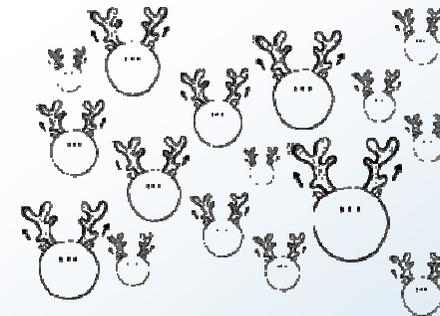
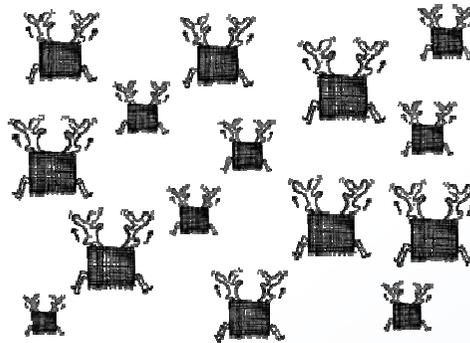
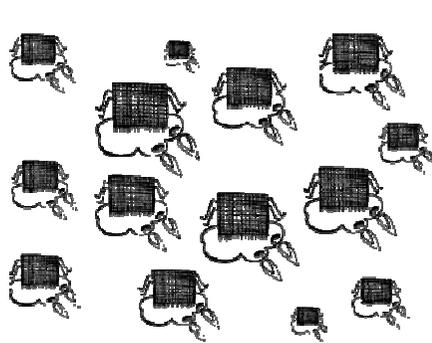
seg bo



fuv bi

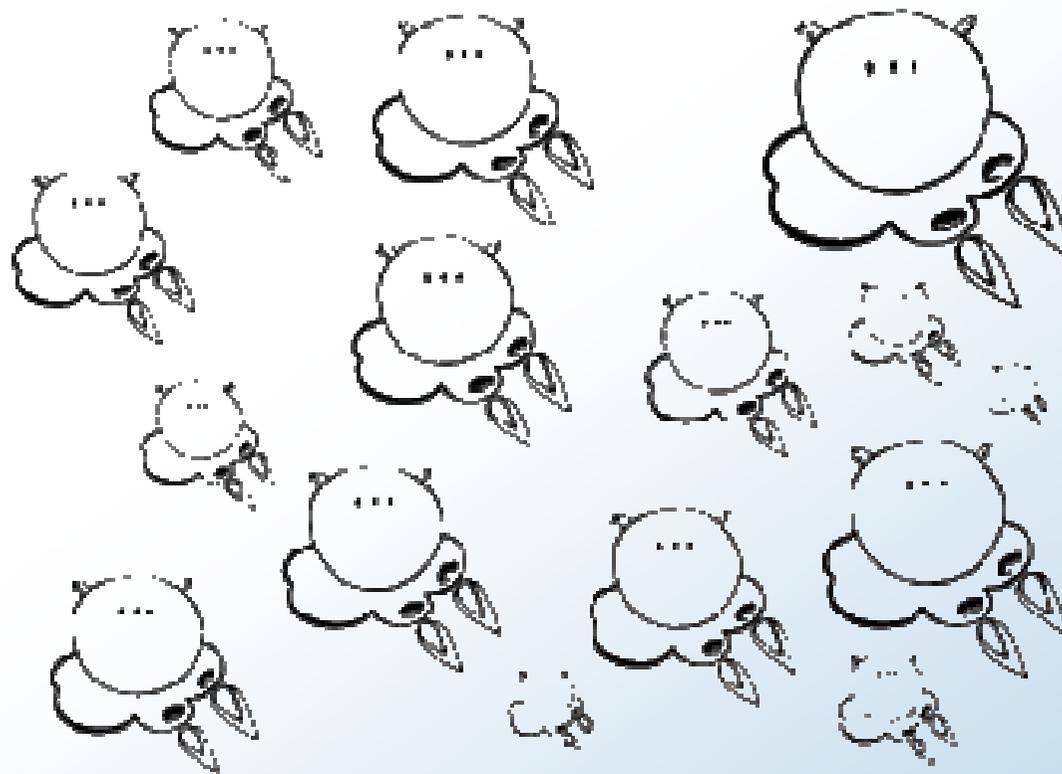
И ТАК ЕЩЕ ДОЛГО

Сеуссе произнес фразу “seg1 bo”. Что он имеет в виду?



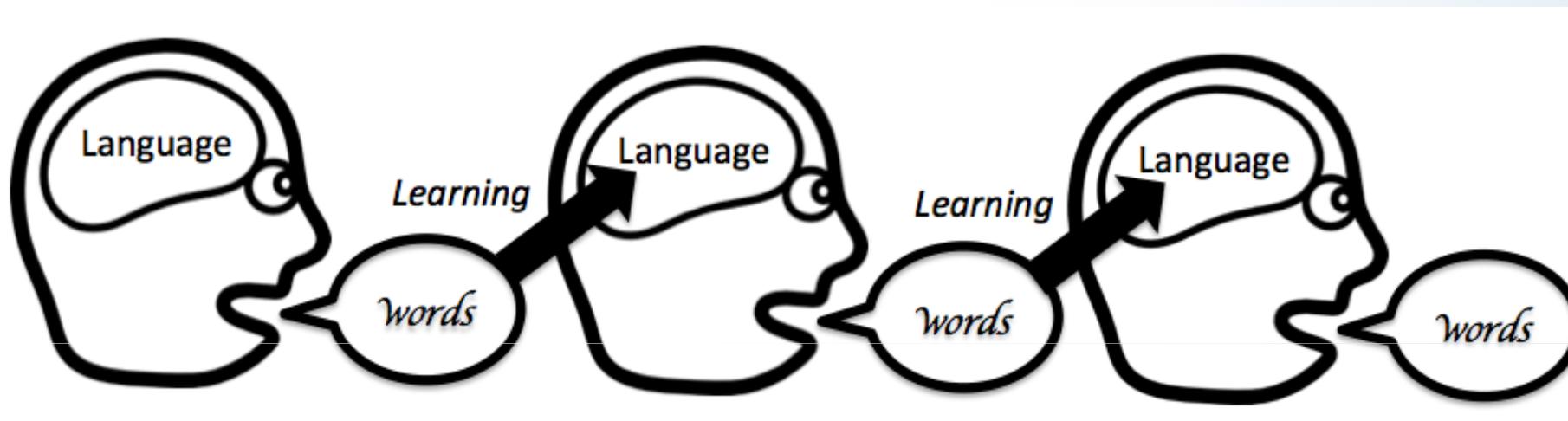
И ТАК ЕЩЕ ДОЛГО

Назовите следующую картинку по-эпсилонски так,
чтобы Сеуссе вас понял



И ТАК ЕЩЕ ДОЛГО

Модель итерированного обучения



ТЕОРИЯ И ИССЛЕДОВАТЕЛЬСКАЯ ГИПОТЕЗА

Почему языки упрощаются или усложняются?

Почему языки упрощаются или усложняются?

Несколько теорий (социолингвистическая типология Традгила, гипотеза языковых ниш Лупяна и Дейла и т.п.) предполагают, что на морфологическую сложность языка могут влиять социальные факторы

Почему языки упрощаются или усложняются?

Несколько теорий (социолингвистическая типология Традгила, гипотеза языковых ниш Лупяна и Дейла и т.п.) предполагают, что на морфологическую сложность языка могут влиять социальные факторы

**«Нормальная» передача языка
из поколения в поколение**

**«Нарушенная» передача языка
из поколения в поколение**

Почему языки упрощаются или усложняются?

Несколько теорий (социолингвистическая типология Традгила, гипотеза языковых ниш Лупяна и Дейла и т.п.) предполагают, что на морфологическую сложность языка могут влиять социальные факторы

«Нормальная» передача

«Нарушенная» передача

Почему языки упрощаются или усложняются?

Несколько теорий (социолингвистическая типология Традгила, гипотеза языковых ниш Лупяна и Дейла и т.п.) предполагают, что на морфологическую сложность языка могут влиять социальные факторы

«Нормальная» передача

«Нарушенная» передача

- Большая доля не-носителей в популяции
- Большая популяция
- Слабые социальные связи
- Низкая социальная стабильность

Почему языки упрощаются или усложняются?

Несколько теорий (социолингвистическая типология Традгила, гипотеза языковых ниш Лупяна и Дейла и т.п.) предполагают, что на морфологическую сложность языка могут влиять социальные факторы

«Нормальная» передача

- Малая доля не-носителей в популяции
- Небольшая популяция
- Тесные социальные связи
- Высокая социальная стабильность

«Нарушенная» передача

- Большая доля не-носителей в популяции
- Большая популяция
- Слабые социальные связи
- Низкая социальная стабильность

Почему языки упрощаются или усложняются?

Несколько теорий (социолингвистическая типология Традгила, гипотеза языковых ниш Лупяна и Дейла и т.п.) предполагают, что на морфологическую сложность языка могут влиять социальные факторы

«Нормальная» передача

- Малая доля не-носителей в популяции
- Небольшая популяция
- Тесные социальные связи
- Высокая социальная стабильность

Способствует сохранению и накапливанию сложности

«Нарушенная» передача

- Большая доля не-носителей в популяции
- Большая популяция
- Слабые социальные связи
- Низкая социальная стабильность

Способствует потере сложности

Почему языки упрощаются или усложняются?

Несколько теорий (социолингвистическая типология Традгила, гипотеза языковых ниш Лупяна и Дейла и т.п.) предполагают, что на морфологическую сложность языка могут влиять социальные факторы

«Нормальная» передача

- Малая доля не-носителей в популяции
- Небольшая популяция
- Тесные социальные связи
- Высокая социальная стабильность

Способствует сохранению и накапливанию сложности

«Нарушенная» передача

- Большая доля не-носителей в популяции
- Большая популяция
- Слабые социальные связи
- Низкая социальная стабильность

Способствует потере сложности

Почему языки упрощаются или усложняются?

Несколько теорий (социолингвистическая типология Традгила, гипотеза языковых ниш Лупяна и Дейла и т.п.) предполагают, что на морфологическую сложность языка могут влиять социальные факторы

«Нормальная» передача

- Малая доля не-носителей в популяции
- Небольшая популяция
- Тесные социальные связи
- Высокая социальная стабильность

Способствует накоплению сложности

«Нарушенная» передача

- Большая доля не-носителей в популяции
- Большая популяция
- Слабые социальные связи
- Низкая социальная стабильность

Способствует потере сложности

На что опираются эти гипотезы?

- Качественный анализ, теоретические объяснения (Trudgill 2011, McWhorter 2001, Wray and Grace 2007, Dahl 2004)

На что опираются эти гипотезы?

- Качественный анализ, теоретические объяснения (Trudgill 2011, McWhorter 2001, Wray and Grace 2007, Dahl 2004)
- Статистические исследования (Luryan and Dale 2010, Bentz and Winter 2013, Szmrecsányi and Kortmann 2008, Kusters 2003)

На что опираются эти гипотезы?

- Качественный анализ, теоретические объяснения (Trudgill 2011, McWhorter 2001, Wray and Grace 2007, Dahl 2004)
- Статистические исследования (Luryan and Dale 2010, Bentz and Winter 2013, Szmrecsányi and Kortmann 2008, Kusters 2003)
- Диахронические исследования (Carroll et al. 2012)

На что опираются эти гипотезы?

- Качественный анализ, теоретические объяснения (Trudgill 2011, McWhorter 2001, Wray and Grace 2007, Dahl 2004)
- Статистические исследования (Luryan and Dale 2010, Bentz and Winter 2013, Szmrecsányi and Kortmann 2008, Kusters 2003)
- Диахронические исследования (Carroll et al. 2012)
- Компьютерные модели (Real, Chater and Christiansen 2014)

На что опираются эти гипотезы?

- Качественный анализ, теоретические объяснения (Trudgill 2011, McWhorter 2001, Wray and Grace 2007, Dahl 2004)
- Статистические исследования (Luryan and Dale 2010, Bentz and Winter 2013, Szmrecsányi and Kortmann 2008, Kusters 2003)
- Диахронические исследования (Carroll et al. 2012)
- Компьютерные модели (Realí, Chater and Christiansen 2014)
- [По-русски: Бердичевский 2012; Вахтин 2014; Даль 2009]

На что опираются эти гипотезы?

- Качественный анализ, теоретические объяснения (Trudgill 2011, McWhorter 2001, Wray and Grace 2007, Dahl 2004)
- Статистические исследования (Luryan and Dale 2010, Bentz and Winter 2013, Szmrecsányi and Kortmann 2008, Kusters 2003)
- Диахронические исследования (Carroll et al. 2012)
- Компьютерные модели (Realí, Chater and Christiansen 2014)
- [По-русски: Бердичевский 2012; Вахтин 2014; Даль 2009]

+ЭКСПЕРИМЕНТ!

ДИЗАЙН ГЛАЗАМИ ИССЛЕДОВАТЕЛЯ

Участники

- Рекрутировались с помощью объявлений в научно-популярных онлайн-СМИ («Элементы», «Чердак»))
- Взрослые (>16), нелингвисты, носители русского

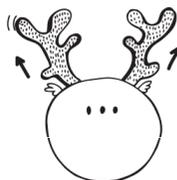
ЯЗЫКИ



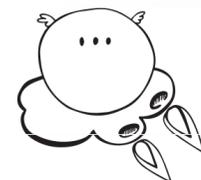
seg



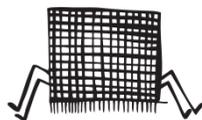
seg mo



seg ro



seg bo



fuv



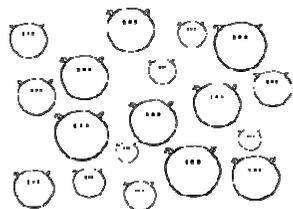
fuv mi



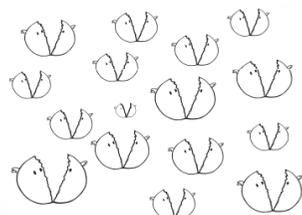
fuv ri



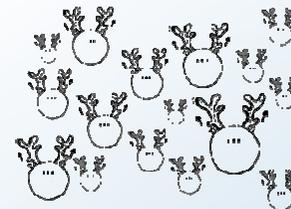
fuv bi



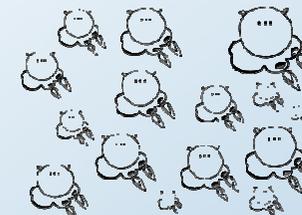
segl



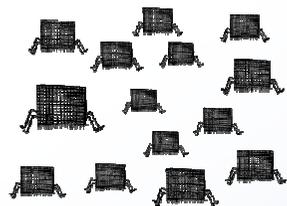
segl mo



segl ro



segl bo



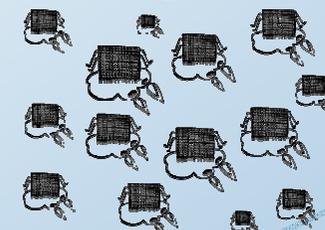
fuvl



fuvl mi



fuvl ri



fuvl bi

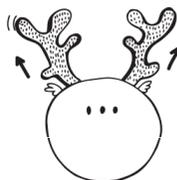
ЯЗЫКИ



seg



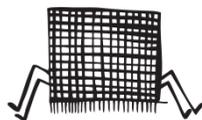
seg m-o



seg r-o



seg b-o



fuv



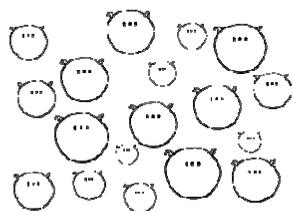
fuv m-i



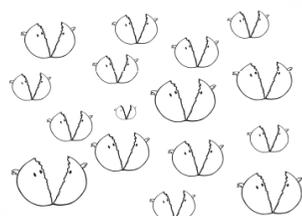
fuv r-i



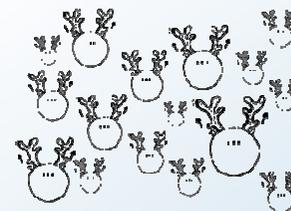
fuv b-i



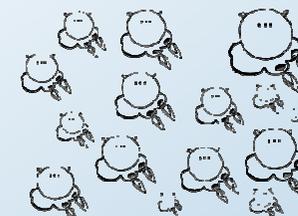
seg-l



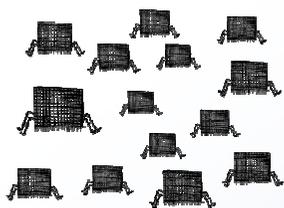
seg-l m-o



seg-l r-o



seg-l b-o



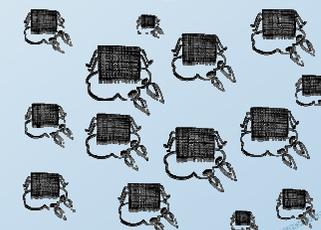
fuv-l



fuv-l m-i



fuv-l r-i



fuv-l b-i

Структура популяций

- Носитель = 6 раундов обучения (по 16 стимулов) + 5 тренировок
- Не-носитель = 3 раунда обучения (по 16 стимулов) + 2 тренировки

	0	1	2	3	4	5	6	7	8	9	10
Норм.	I_n	L1									
Наруш.	I_n	L1	L2	L2	L2	L1	L1	L1	L1	L1	L1

Структура популяций

- Носитель = 6 раундов обучения (по 16 стимулов) + 5 тренировок
- Не-носитель = 3 раунда обучения (по 16 стимулов) + 2 тренировки

	0	1	2	3	4	5	6	7	8	9	10
Норм.	I_n	L1									
Наруш.	I_n	L1	L2	L2	L2	L1	L1	L1	L1	L1	L1



КОНТАКТ

Структура популяций

- Носитель = 6 раундов обучения (по 16 стимулов) + 5 тренировок
- Не-носитель = 3 раунда обучения (по 16 стимулов) + 2 тренировки

	0	1	2	3	4	5	6	7	8	9	10
Норм.	I_n	L1									
Наруш.	I_n	L1	L2	L2	L2	L1	L1	L1	L1	L1	L1



КОНТАКТ

10 нормальных цепочек, 10 «нарушенных» цепочек (на тех же 10 языках)

Структура популяций

- Носитель = 6 раундов обучения (по 16 стимулов) + 5 тренировок
- Не-носитель = 3 раунда обучения (по 16 стимулов) + 2 тренировки

	0	1	2	3	4	5	6	7	8	9	10
Норм.	I_n	L1									
Наруш.	I_n	L1	L2	L2	L2	L1	L1	L1	L1	L1	L1



КОНТАКТ

10 нормальных цепочек, 10 «нарушенных» цепочек (на тех же 10 языках)

Структура популяций

- Носитель = 6 раундов обучения (по 16 стимулов) + 5 тренировок
- Не-носитель = 3 раунда обучения (по 16 стимулов) + 2 тренировки

	0	1	2	3	4	5	6	7	8	9	10
Норм.	I_n	L1									
Наруш.	I_n	L1	L2	L2	L2	L1	L1	L1	L1	L1	L1



КОНТАКТ

10 нормальных цепочек, 10 «нарушенных» цепочек (на тех же 10 языках)
Предсказание: в нормальных цепочках язык последнего поколения а) сохранит род с большей вероятностью б) в среднем будет сложнее, чем в нарушенных цепочках. На выразительности языков это не скажется.

Структура популяций

- Носитель = 6 раундов обучения (по 16 стимулов) + 5 тренировок
- Не-носитель = 3 раунда обучения (по 16 стимулов) + 2 тренировки

	0	1	2	3	4	5	6	7	8	9	10
Норм.	I_n	L1									
Наруш.	I_n	L1	L2	L2	L2	L1	L1	L1	L1	L1	L1



КОНТАКТ

10 нормальных цепочек, 10 «нарушенных» цепочек (на тех же 10 языках)
Предсказание: в нормальных цепочках язык последнего поколения а) сохранит *род* с большей вероятностью б) в среднем будет сложнее, чем в нарушенных цепочках. На выразительности языков это не скажется.

РЕЗУЛЬТАТЫ

Беглый качественный анализ (нормальные цепочки)

- Две цепочки сохранили морфологию и даже увеличили сложность
- Одна цепочка утратила род, но стала согласовывать глаголы по числу
- Одна цепочка утратила род
- Остальные шесть в двух словах описываются плохо

Увеличение сложности-1

		exist	break	antlers	fly
Animal 1	sg	rub	rub le	rub ze	rub ne
Animal 2	sg	vad	vad lo	vad zo	vad no
Animal 1	pl	rub s	rub s le	rubp ze	rubp ne
Animal 2	pl	vadp	vadp lo	vadp zo	vadp no

Увеличение сложности-2

		exist	break	antlers	fly
Animal 1	sg	jag	jag se	jag ke	jag de
Animal 2	sg	lor	lor su	lor ku	lor du
Animal 1	pl	jagv	jagv se	jagv e	jagv de
Animal 2	pl	lorv	lorv su	lorv ku	lorv du

Согласование по числу

		exist	break	antlers	fly
Animal 1	sg	sig	sig je	sig ve	sig se
Animal 2	sg	hop	hop je	hop ve	hop se
Animal 1	pl	sig	sigm ju	sigm vu	sigm su
Animal 2	pl	hop	hopm ju	hopm vu	hopm su

Согласование по числу (!)

		exist	break	antlers	fly
Animal 1	sg	sig	sig je	sig ve	sig se
Animal 2	sg	hop	hop je	hop ve	hop se
Animal 1	pl	sig	sigm ju	sigm vu	sigm su
Animal 2	pl	hop	hopm ju	hopm vu	hopm su

Утрата рода

		exist	break	antlers	fly
Animal 1	sg	fim	fim bu	fim jo	fim su
Animal 2	sg	caz	fim bu	caz jo	caz su
Animal 1	pl	fimh	fimh bu	fimh jo	fimh su
Animal 2	pl	cazh	cazh bu	cazh jo	cazh su

Плохо описываемый результат

		exist	break	antlers	fly
Animal 1	sg	dig	dig mo	dig po	dig ho
Animal 2	sg	sez	sez mu	sez po	sez ho
Animal 1	pl	dign	dign mu	dign po	sez
Animal 2	pl	sezn	sezn mu	sezn po	sezn ho

Беглый качественный анализ (нарушенные цепочки)

- Одна утратила род полностью и «чисто»
- Одна утратила род почти полностью
- Две утратили род полностью, но с неожиданным дополнительным упрощением
- Одна утратила род полностью, но перешла к нетривиальному выражению числа
- Остальные пять в двух словах описываются плохо
- Ни одна не сохранила род полностью

Утрата рода и частичная утрата числа

		exist	break	antlers	fly
Animal 1	sg	noj	noj ga	noj ca	noj ce
Animal 2	sg	ful	ful ga	ful ca	ful ce
Animal 1	pl	nojb	noj ga	nojb ca	nojb ce
Animal 2	pl	fulb	ful ga	fulb ca	fulb ce

Утрата рода и лексическая омонимия

		exist	break	antlers	fly
Animal 1	sg	sig	sig ce	sig ju	sig vo
Animal 2	sg	hop	hop	hop ju	hop vo
Animal 1	pl	sigm	sigm ce	sigm ju	sigm vo
Animal 2	pl	hopm	hopm	hopm ju	hopm vo

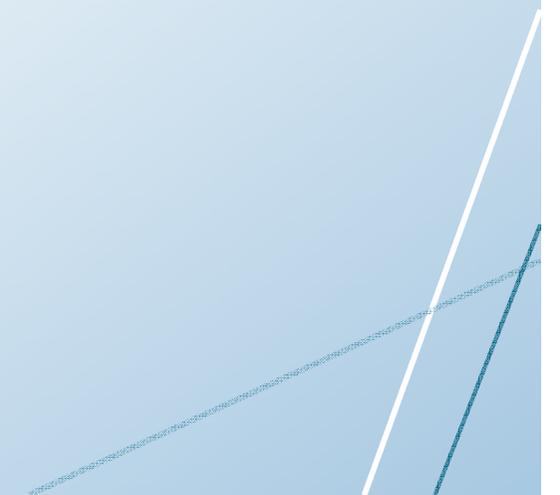
Утрата рода и лексическая омонимия (!)

		exist	break	antlers	fly
Animal 1	sg	sig	sig ce	sig ju	sig vo
Animal 2	sg	hop	hop	hop ju	hop vo
Animal 1	pl	sigm	sigm ce	sigm ju	sigm vo
Animal 2	pl	hopm	hopm	hopm ju	hopm vo

Утрата рода и изысканное выражение числа

		exist	break	antlers	fly
Animal 1	sg	senz	senz po	senz ho	senz mo
Animal 2	sg	sign	sign po	sign ho	sign mo
Animal 1	pl	sezn	sezn po	sezn ho	sezn mo
Animal 2	pl	dign	dign po	dign ho	dign mo

Количественный анализ



Количественный анализ

- **Выразительность языка:** доля правильных ответов в тесте на понимание

Количественный анализ

- Выразительность языка: доля правильных ответов в тесте на понимание
- Скорость изменений: сумма нормализованных расстояний Левенштейна между описаниями одного и того же стимула в языках смежных поколений
- Сложность

Количественный анализ

- Выразительность языка: доля правильных ответов в тесте на понимание
- Скорость изменений: сумма нормализованных расстояний Левенштейна между описаниями одного и того же стимула в языках смежных поколений
- Сложность: RegMap (регулярность соответствий; Tamariz 2011; Cornish, Tamariz and Kirby 2010; Tamariz and Smith 2008)

RegMap

$M = \{animals; events\}$

$S = \{seg, fuv, segl, fuvl, ro, mo, bo, ri, mi, bi\}$

RegMap

$M = \{animals; events\}$

$S = \{seg, fuv, segl, fuvl, ro, mo, bo, ri, mi, bi\}$

$$RegMap(S|M) = 1 - \frac{H(S|M)}{\log(n_s)}$$

RegMap

$M = \{animals; events\}$

$S = \{seg, fuv, segl, fuvl, ro, mo, bo, ri, mi, bi\}$

$$RegMap(S|M) = 1 - \frac{H(S|M)}{\log(n_s)}$$

$$RegMap(M|S) = 1 - \frac{H(M|S)}{\log(n_m)}$$

RegMap

$M = \{animals; events\}$

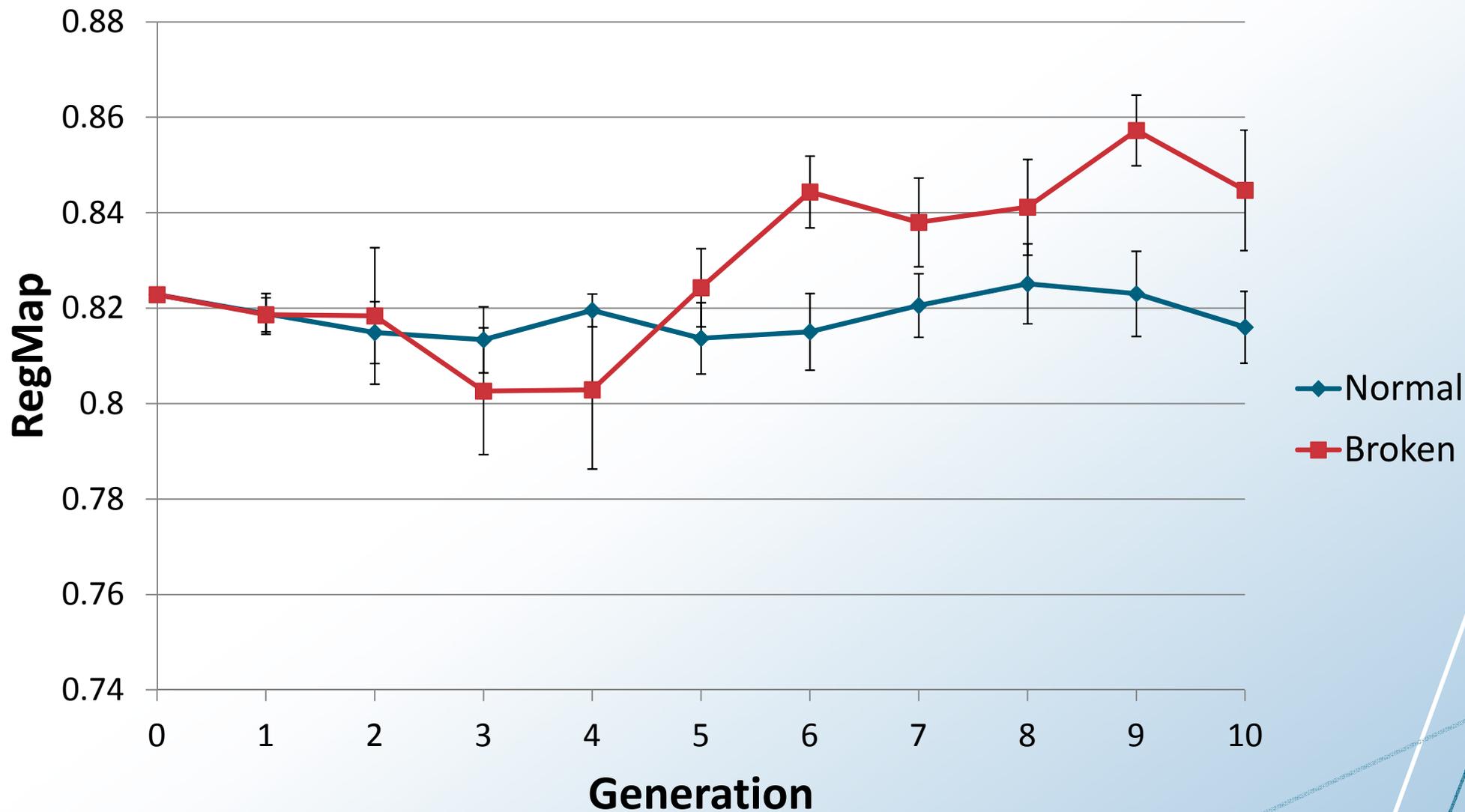
$S = \{seg, fuv, segl, fuvl, ro, mo, bo, ri, mi, bi\}$

$$RegMap(S|M) = 1 - \frac{H(S|M)}{\log(n_s)}$$

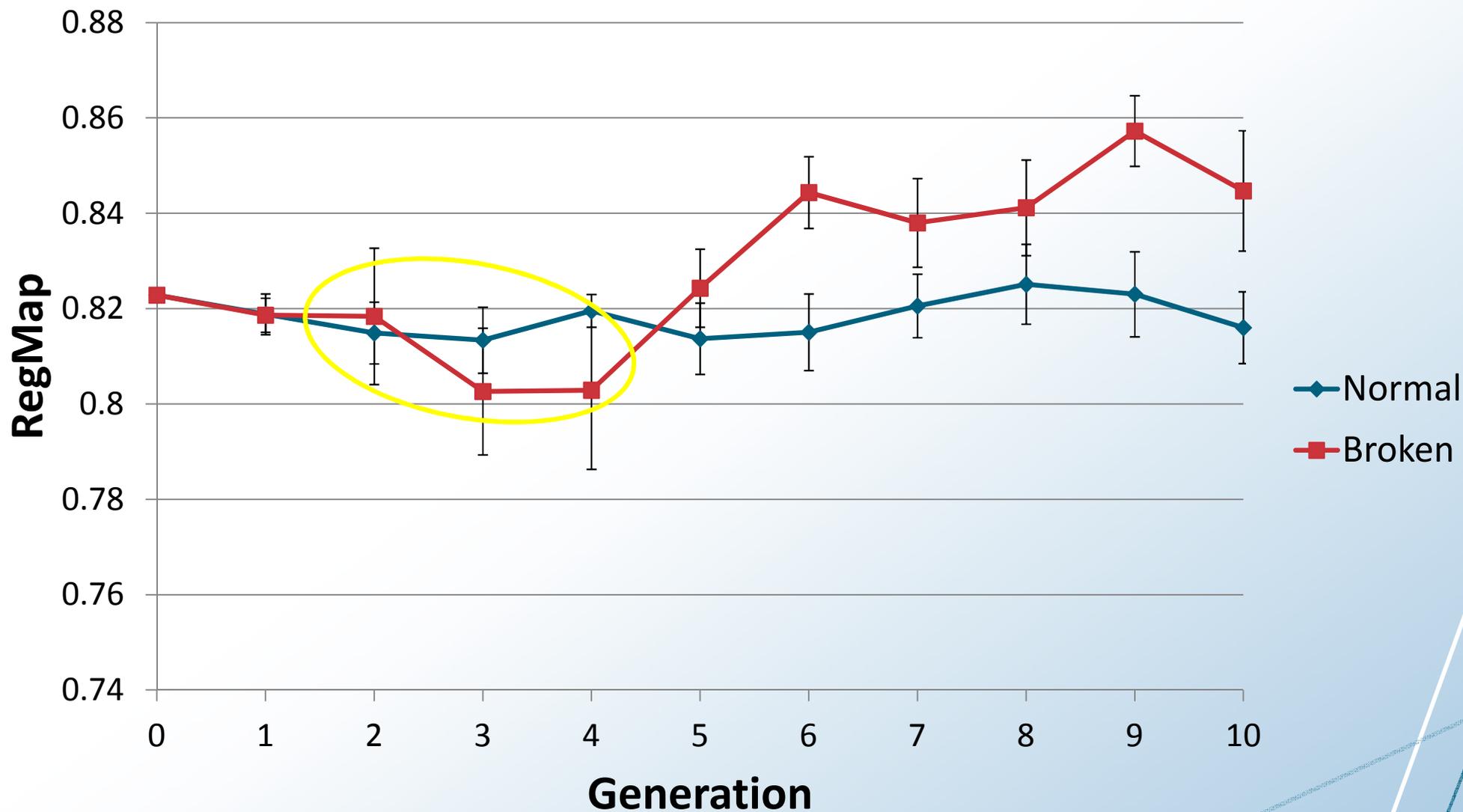
$$RegMap(M|S) = 1 - \frac{H(M|S)}{\log(n_m)}$$

$$RegMap = 2 * \frac{RegMap(S|M) * RegMap(M|S)}{RegMap(S|M) + RegMap(M|S)}$$

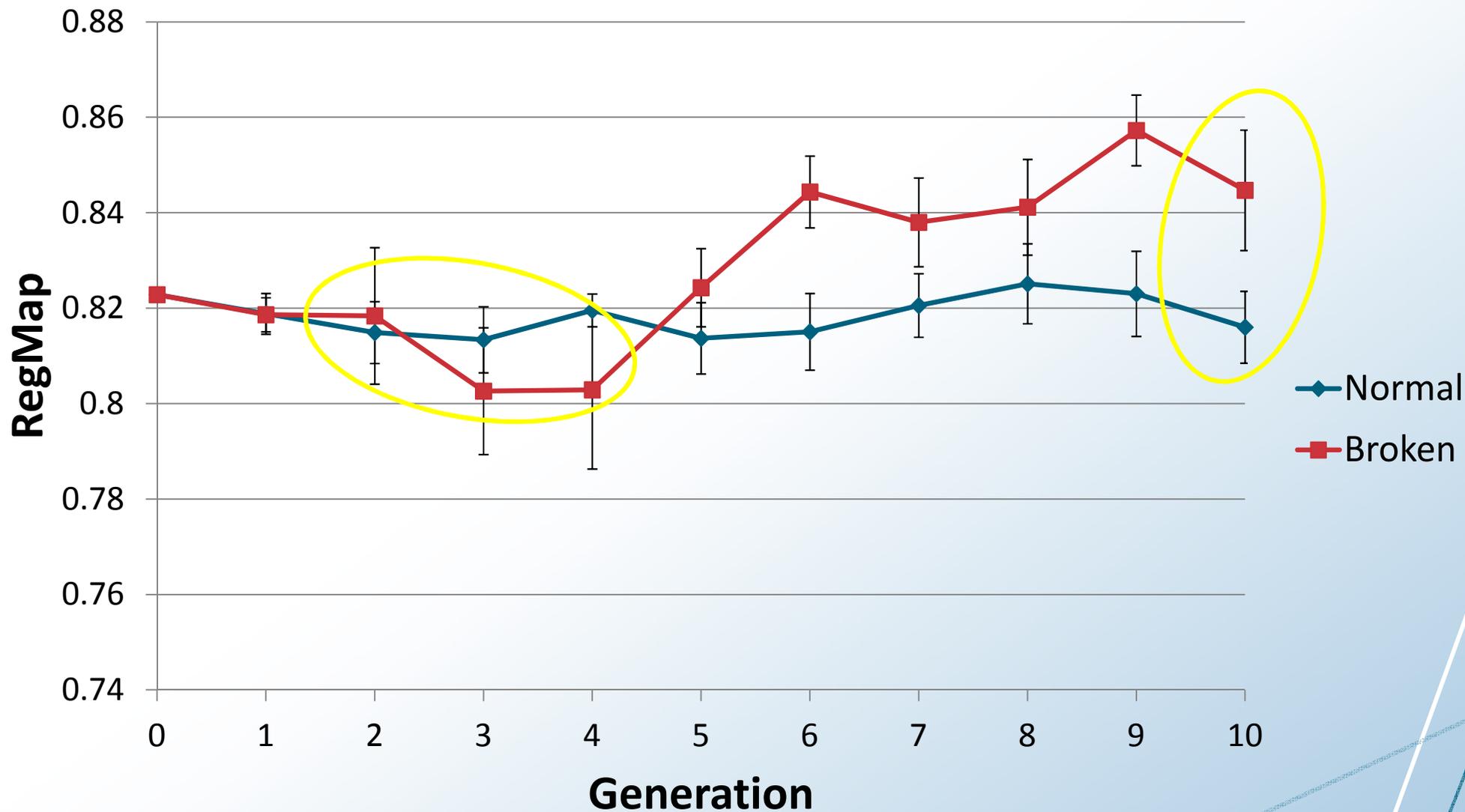
Изменение сложности



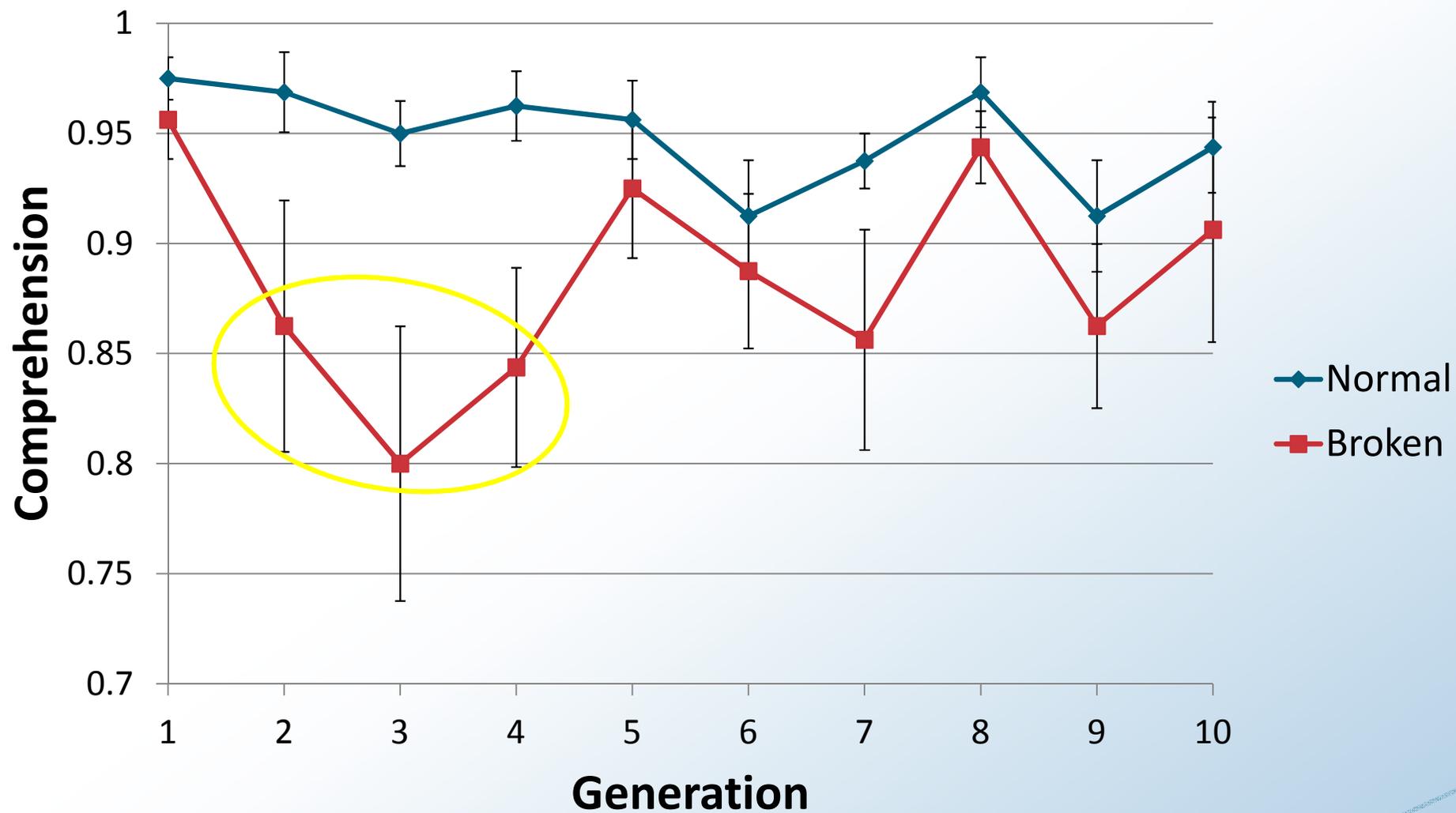
Изменение сложности



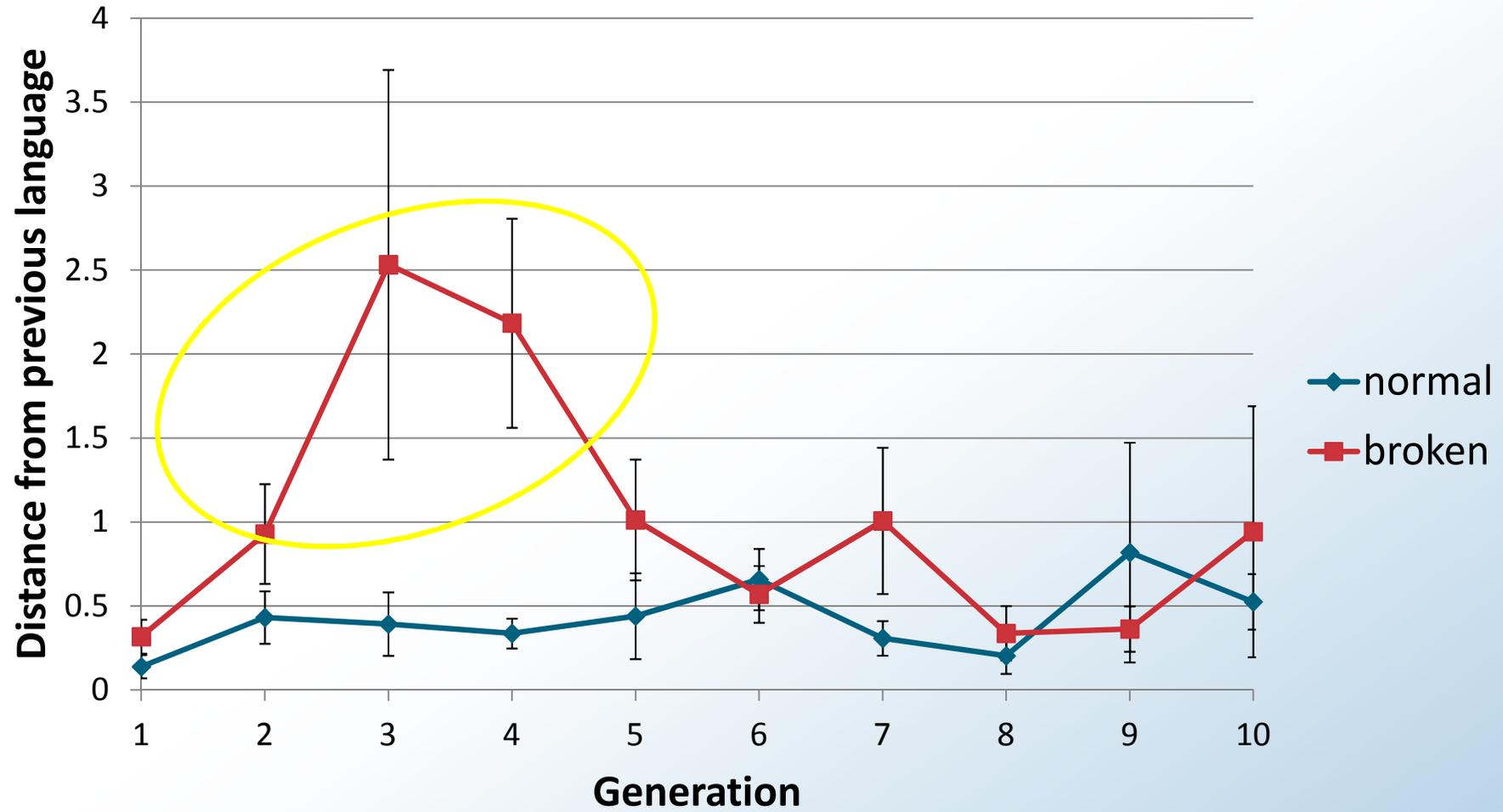
Изменение сложности



Изменение выразительности



Скорость изменений



Почему же в «контактный» период наблюдается усложнение?

- Неполное освоение грамматических категорий не-носителями приводит не к обобщению, а к свободному варьированию.

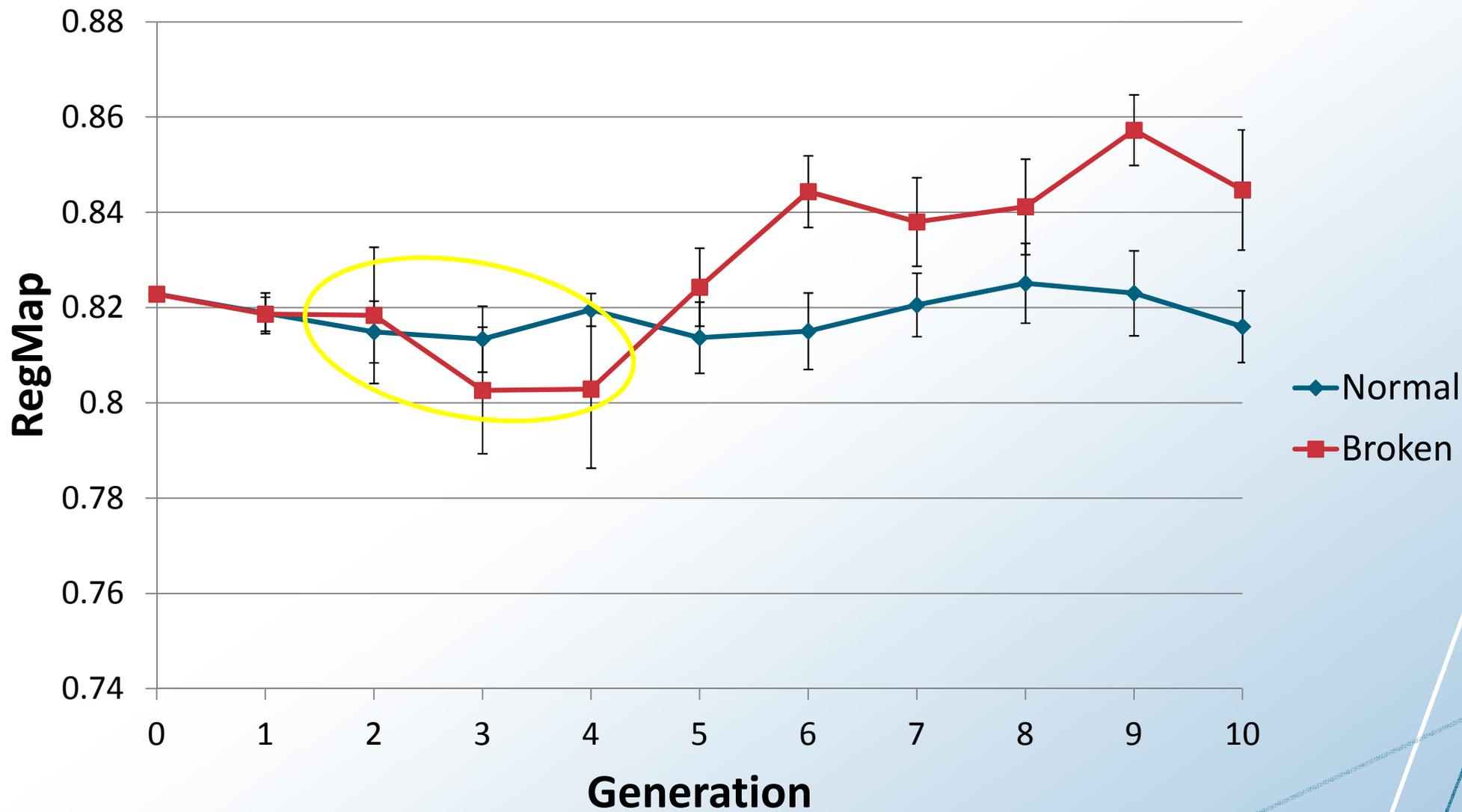
Почему же в «контактный» период наблюдается усложнение?

- Неполное освоение грамматических категорий не-носителями приводит не к обобщению, а к свободному варьированию.
- Это уничтожает морфологию, но регулярность соответствий не повышается, а понижается (к тому же добавляются другие типы немотивированной вариативности)

Почему же в «контактный» период наблюдается усложнение?

- Неполное освоение грамматических категорий не-носителями приводит не к обобщению, а к свободному варьированию.
- Это уничтожает морфологию, но регулярность соответствий не повышается, а понижается
- После контакта носители избавляются от непредсказуемой вариативности, что приводит к наблюдаемому упрощению

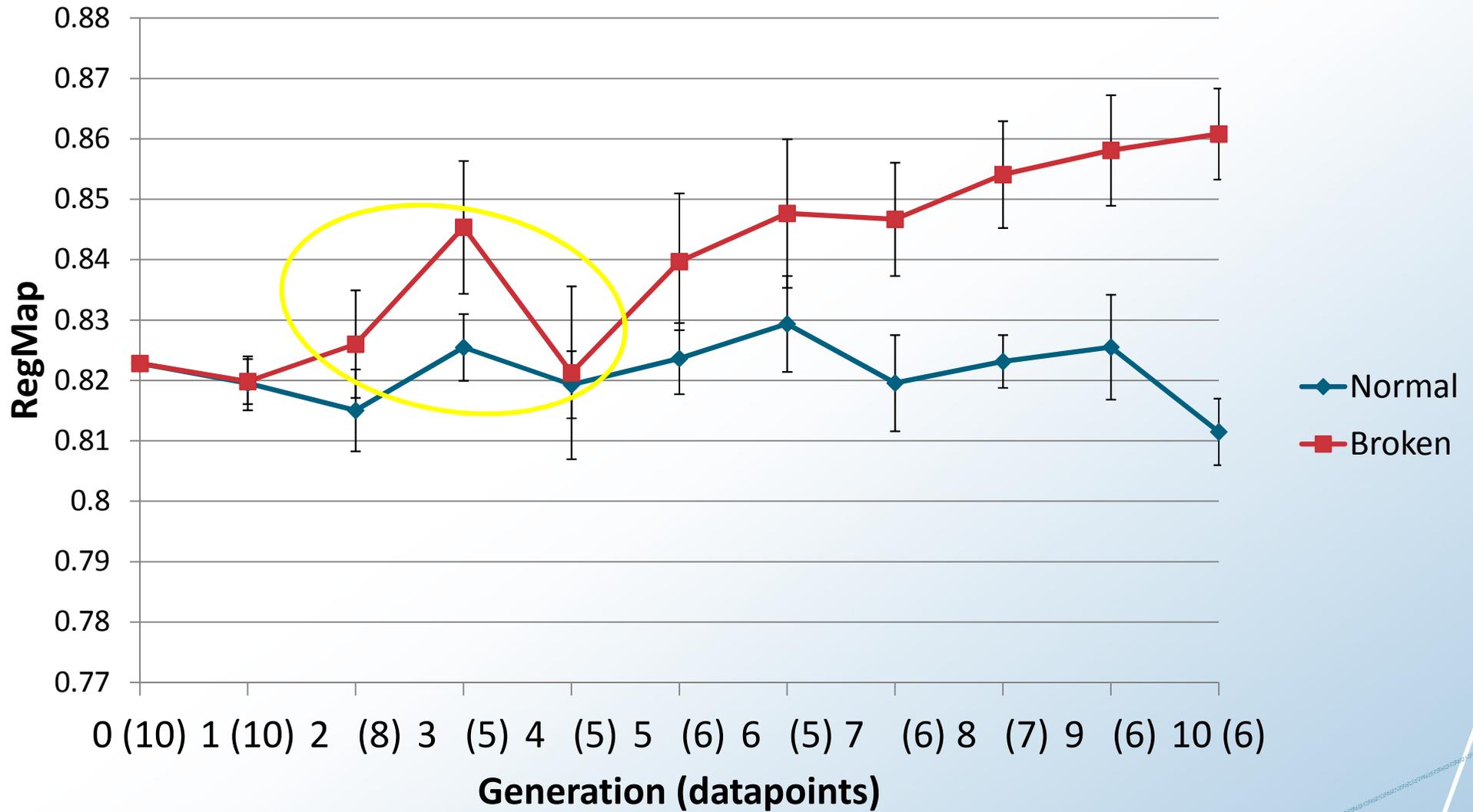
Изменение сложности



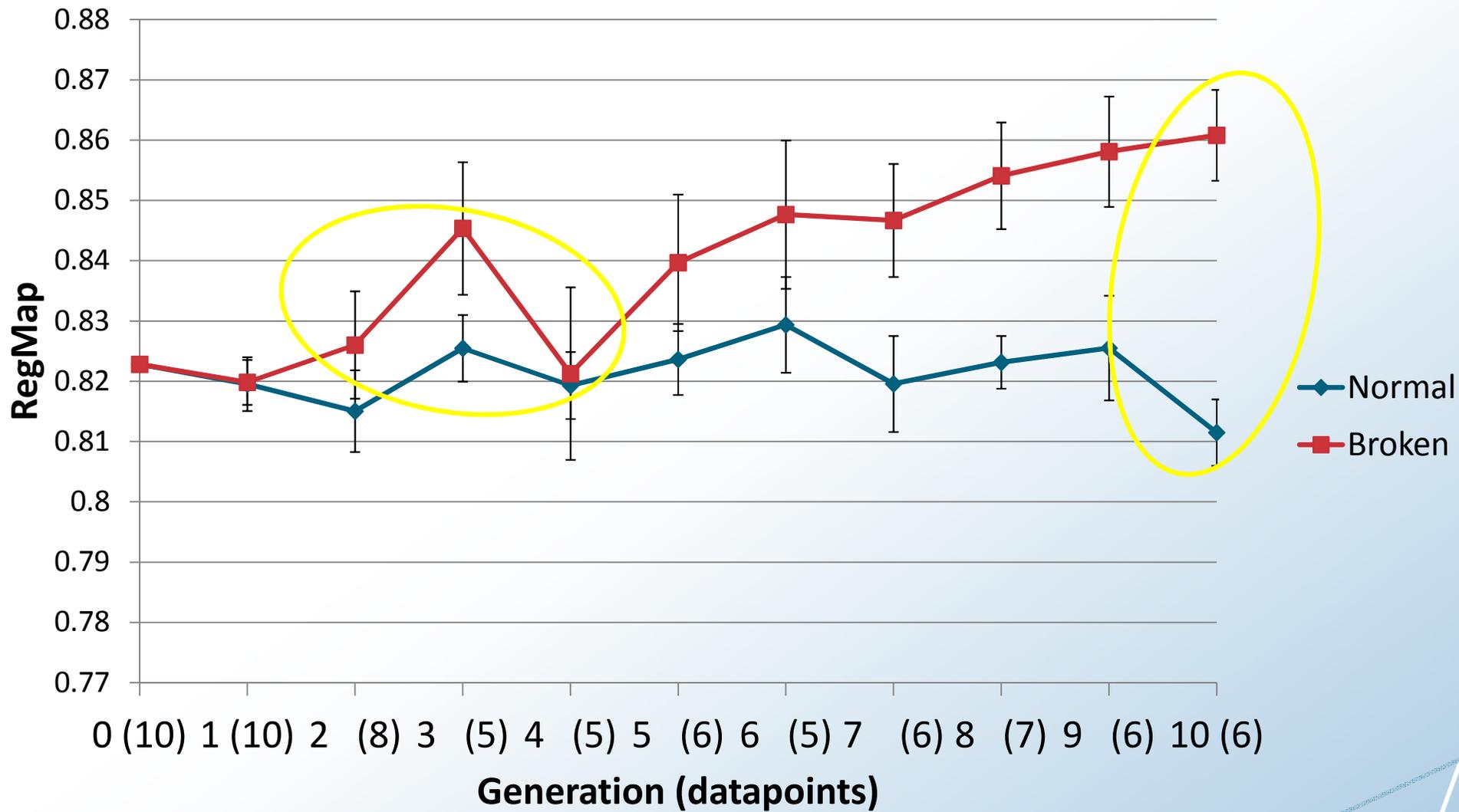
Почему же в «контактный» период наблюдается усложнение?

- Неполное освоение грамматических категорий не-носителями приводит не к обобщению, а к свободному варьированию.
- Это уничтожает морфологию, но регулярность соответствий не повышается, а понижается
- После контакта носители избавляются от непредсказуемой вариативности, что приводит к наблюдаемому упрощению
- Сделаем поправку на нестабильность: удалим все знаки, которые участник не распознал в тесте на понимание

Изменение сложности (с поправкой на нестабильность)



Изменение сложности (с поправкой на нестабильность)



Заключение

- Предсказание выполняется: итоговые языки в нарушенных цепочках оказались проще, чем в нормальных

Заключение

- Предсказание выполняется: итоговые языки в нарушенных цепочках оказались проще, чем в нормальных
- Разрушение морфологии в нормальных цепочках -- (сравнительно безвредный) методологический артефакт

Заключение

- Предсказание выполняется: итоговые языки в нарушенных цепочках оказались проще, чем в нормальных
- Разрушение морфологии в нормальных цепочках -- (сравнительно безвредный) методологический артефакт
- Усложнение в контактный период -- вероятно, следствие нестабильности ментальных грамматик говорящих, приводящей к непредсказуемой вариативности (ср. иную картину при поправке на нестабильность)

Открытые вопросы

- Усложнение за счет непредсказуемой вариативности: наблюдается ли этот эффект у не-носителей в реальности?
- Настолько ли он силен?
- Настолько ли велика его роль в популяции?
- Действительно ли механизм упрощения за счет контакта устроен именно так?
- Насколько адекватна модель?

Благодарности

- Елене Мартыновой и «Элементам» (elementy.ru), Ире Якутенко и «Чердаку» (chrdk.ru) за рекламу
- Тане Руссите (tanjarussita.com) за эпсилонскую фауну
- Исследовательской группе CLEAR за финансирование
- Бета-тестерам за первое десантирование на Эпсилон

