WHAT IS A POSSIBLE WORD? EVIDENCE FROM RUSSIAN FACTITIVE VERBS

Anna Endresen

Laura A. Janda

CLEAR (Cognitive Linguistics: Empirical Approaches to Russian)

University of Tromsø

What happens in the gap between the actual and the impossible in a language?

ACTUAL

IMPOSSIBLE

What happens in the gap between the actual and the impossible in a language?

ACTUAL

POSSIBLE

IMPOSSIBLE

What happens in the gap between the actual and the impossible in a language?

ACTUAL

STANDARD WORDS

Words that are standard and conventionalized, might be stored in memory rather than generated on the fly

POSSIBLE

MARGINAL WORDS

Words that are generated by some speakers and can be understood / accepted by some speakers

IMPOSSIBLE

NONCE WORDS

Words that cannot be generated and do not exist

What happens in the gap between the actual and the impossible in a language?

ACTUAL

STANDARD WORDS

Words that are standard and conventionalized, might be stored in memory rather than generated on the fly

e.g.: *undo*

POSSIBLE

MARGINAL WORDS

Words that are generated by some speakers and can be understood / accepted by some speakers

e.g.: *unworry*

IMPOSSIBLE

NONCE WORDS

Words that cannot be generated and do not exist

e.g.: *unblick*

What is a possible (marginal) word?

A word which
- is attested at least once
- is not established in standard language
- is a spontaneous creation generated on the fly, on a certain occasion
- is generated on the basis of a productive morphological pattern
- is analyzable and semantically transparent

- **It exists and at the same time it does not exist.**

Case study: Russian factitive verbs

- *Ob''jasnit'* 'clarify, make X be clear' < *jasnyj* 'clear$_{ADJ}$'

- **Two most productive patterns**: o-…-it' and u-…-it'

- We are interested in **new coinages** like
  - *omuzykalit'* 'musicalize' (< *muzykal'nyj* 'musical$_{ADJ}$')
  - *ukonkretit'* 'concretize' (< *konkretnyj* 'concrete$_{ADJ}$')
  - *ovnešnit'* 'externalize' (< *vnešnij* 'external$_{ADJ}$')

Productivity: competence or performance?

- Usually newly coined words are examined from the perspective of linguistic performance (Haspelmath 2002: 112).

- "A widespread view among linguists is that linguistic competence and linguistic performance are conceptually quite distinct and should therefore be studied separately. **The different degrees of productivity that we observe in word-formation are a problem** for this view, because **rule productivity is not clearly a property of either competence or performance.**" (Haspelmath 2002: 110)

- We offer a study of productivity that combines both perspectives: performance and competence. **Frequency** reflects **performance**, whereas knowledge of **productive patterns** reflects **competence**.

Performance:
We look at novel marginal factitive verbs generated by some native speakers and attested in the corpus
Competence:
We want to know how these possible words are perceived by other native speakers?

Hypothesis 1:	Standard		Marginal		Nonce
Hypothesis 2:	Standard		Marginal		Nonce
Hypothesis 3:	Standard		Marginal		Nonce
Hypothesis 4:	Standard	gradual transition	Marginal	gradual transition	Nonce

Experimental design

STANDARD WORDS

Words that are standard and conventionalized, might be stored in memory rather than generated on the fly

e.g.: *undo*

MARGINAL WORDS

Words that are generated by some speakers and can be understood / accepted by some speakers

e.g.: *unworry*

NONCE WORDS

Words that cannot be generated and do not exist

e.g.: *unblick*

Experimental design

STANDARD WORDS	MARGINAL WORDS	NONCE WORDS
10 o + 10 u	10 o + 10 u	10 o + 10 u

ob''jasnit' *uskorit'*	*oser'ëznit'* *uvkusnit'*	*osurit'* *usaglit'*

60 stimuli

- All factitives used in the experiment are **deadjectival**. This decision is made in order to reduce the number of valuables.

- All standard and marginal factitives chosen for experiment are **morphologically transparent** and **analyzable** and have a clear existing adjectival base.

- All factitives are given as **perfective infinitives**.

- Factitive verbs are presented **in contexts**.
  - For standard and marginal factitives we are using real contexts from the corpus, often shortened.
  - The contexts of nonce factitive verbs are based on corpus contexts of real verbs with meanings similar to those that are assumed for nonce factitives.

Standard factitives

- All chosen standard factitives are highly frequent in the RNC.

#	O- factitive	Freq RNC 1950-2012	U- factitive	Freq RNC 1950-2012
1	*ob''jasnit'*	18,149	*utočnit'*	2,860
2	*oblegčit'*	1,802	*umen'šit'*	2,010
3	*oslabit'*	1,401	*uskorit'*	2,008
4	*okruglit'*	939	*ulučšit'*	1,899
5	*obogatit'*	800	*uprostit'*	1,350
6	*ožestočit'*	686	*ukorotit'*	787
7	*osložnit'*	410	*usložnit'*	311
8	*ogolit'*	387	*uteplit'*	205
9	*osčastlivit'*	343	*uplotnit'*	201
10	*osvežit'*	280	*uxudšit'*	199

Marginal factitives – POSSIBLE WORDS

- All marginal factitives have low token frequency in the RNC

#	O- factitive	Freq RNC 1950-2012	U- factitive	Freq RNC 1950-2012
1	omeždunarodit'	1	uvkusnit'	1
2	opoxabit'	1	umedlit'	1
3	opriličit'	1	ukrasivit'	1
4	oser'eznit'	1	user'eznit'	1
5	ostekljanit'	1	usovremenit'	1
6	oržavit'	2	ukonkretit'	1
7	osurovit'	2	ustrožit'	3
8	obytovit'	3	ucelomudrit'	3
9	ovnešnit'	4	uprozračit'	4
10	omuzykalit'	4	udorožit'	8

Nonce factitives

- O- and U- nonce factitives correlate with respect to the initial consonant of the stem:

- Variety of initial stem consonants

- Nonce words should not be similar to standard or possible factitives so that the speaker should not hesitate about them.

#	O- factitive	U- factitive
1	*osurit'*	*usaglit'*
2	*otovit'*	*utulit'*
3	*oduktit'*	*udamlit'*
4	*ogabit'*	*uguzvit'*
5	*okočlit'*	*ukampit'*
6	*ošaklit'*	*ušadrit'*
7	*očavit'*	*učopit'*
8	*oblusit'*	*uloprit'*
9	*onomit'*	*unokrit'*
10	*obmomlit'*	*umarvit'*

Experiment: score-assignment test

The task: Evaluate the <u>marked</u> <u>word</u> using one of the statements.

Давно пора как-то <u>оприличить</u> наше общение более мягкими выражениями.
'It's high time we <u>made</u> our interaction <u>respectable</u> by using kinder statements.'

☐ 5 points - Это совершенно нормальное слово русского языка.
       'This is an absolutely normal Russian word'

☐ 4 points - Это слово нормальное, но его мало используют.
       'This word is normal, but it is rarely used'

☐ 3 points - Это слово звучит странно, но, может быть, его кто-то использует.
       'This word sounds strange, but someone might use it'

☐ 2 points - Это слово звучит странно, и его вряд ли кто-то использует.
       'This word sounds strange and it is unlikely that anyone uses it'

☐ 1 point - Этого слова в русском языке нет.
       'This word does not exist in the Russian language.'

A few more issues

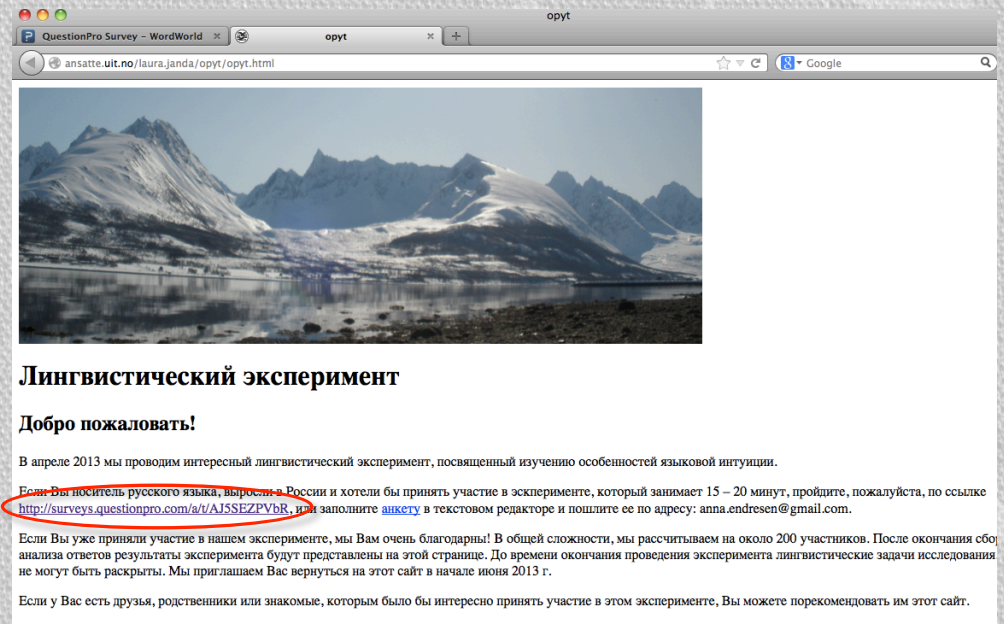## Order of stimuli:
- Semi-random
- The same for all

Administration:
- Questionnaire
- No limits on time

Age groups:
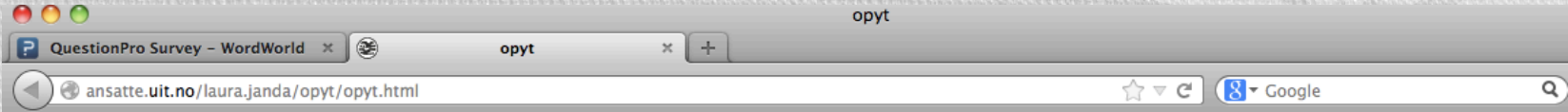- School age children aged 14 – 17: 70 participants
- Adults: 51 participants

Website

http://ansatte.uit.no/laura.janda/opyt/opyt.html

Лингвистический эксперимент

Добро пожаловать!

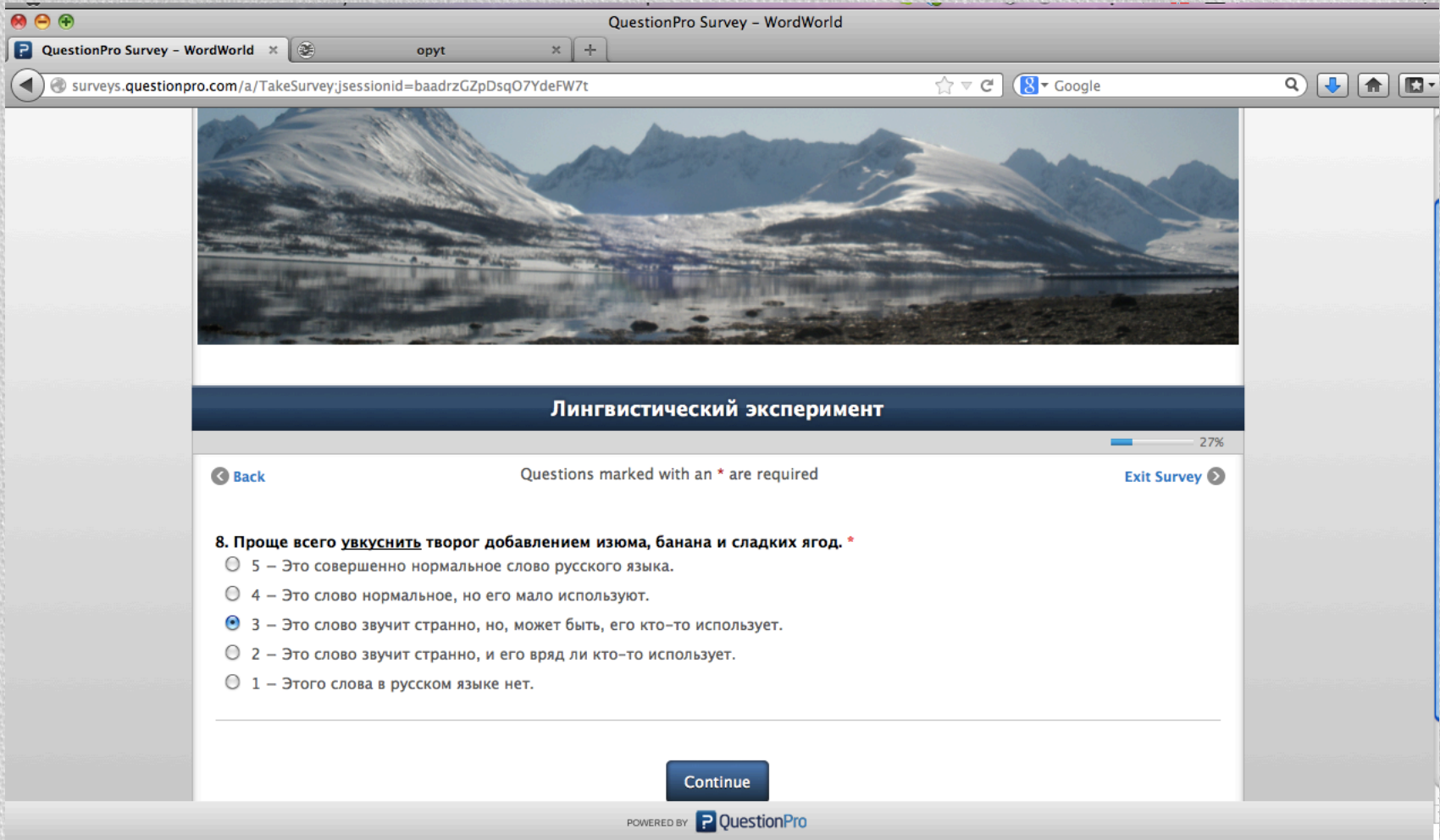В апреле 2013 мы проводим интересный лингвистический эксперимент, посвященный изучению особенностей языковой интуиции.

Если Вы носитель русского языка, выросли в России и хотели бы принять участие в эксперименте, который занимает 15 – 20 минут, пройдите, пожалуйста, по ссылке http://surveys.questionpro.com/a/t/AJ5SEZPVbR, или заполните анкету в текстовом редакторе и пошлите ее по адресу: anna.endresen@gmail.com.

Если Вы уже приняли участие в нашем эксперименте, мы Вам очень благодарны! В общей сложности, мы рассчитываем на около 200 участников. После окончания сбор анализа ответов результаты эксперимента будут представлены на этой странице. До времени окончания проведения эксперимента лингвистические задачи исследования не могут быть раскрыты. Мы приглашаем Вас вернуться на этот сайт в начале июня 2013 г.

Если у Вас есть друзья, родственники или знакомые, которым было бы интересно принять участие в этом эксперименте, Вы можете порекомендовать им этот сайт.

Website
http://ansatte.uit.no/laura.janda/opyt/opyt.html

Questionnaire on-line:
http://surveys.questionpro.com/a/t/AJ5SEZPVbR

A few more issues

Order of stimuli:
- Semi-random
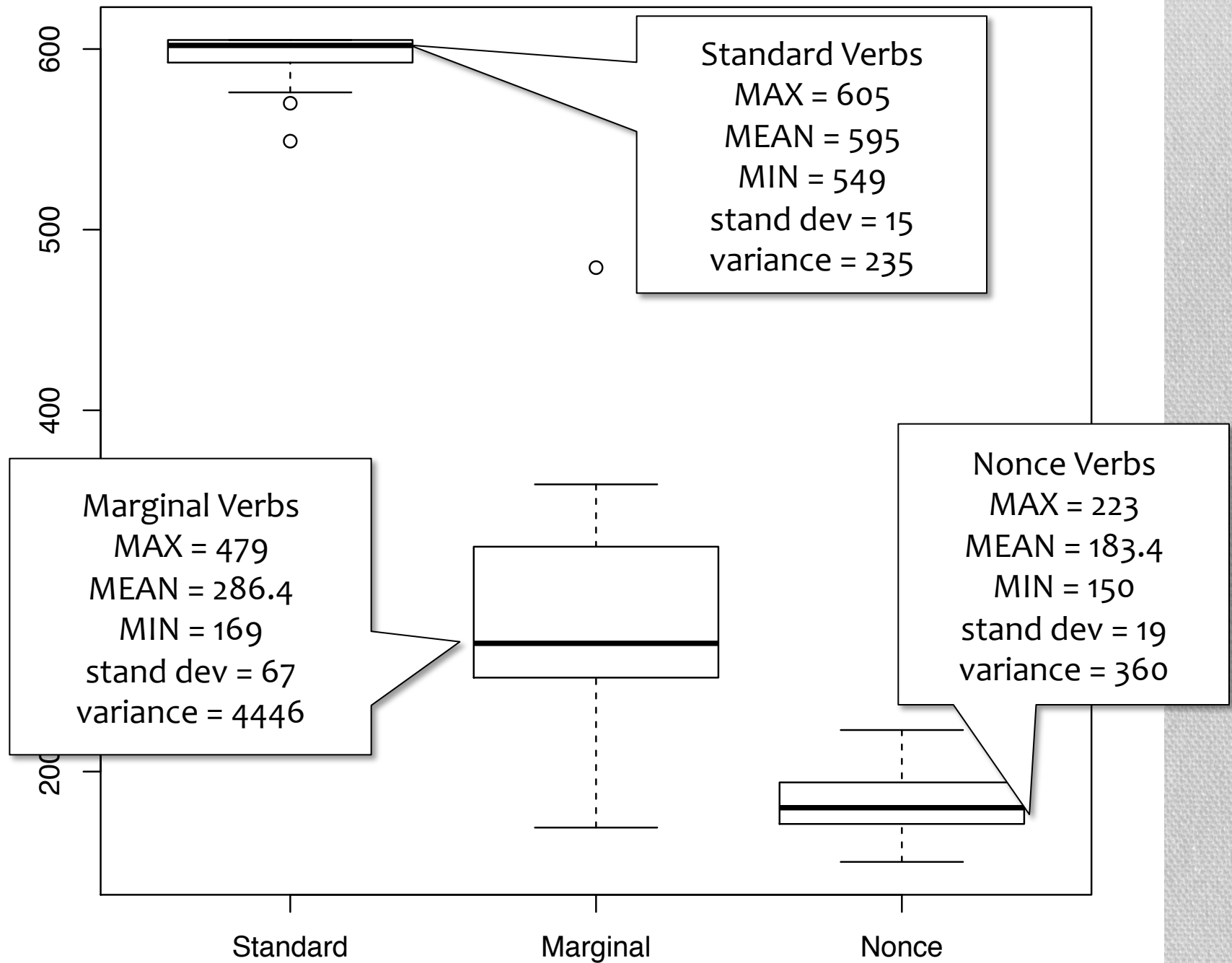- The same for all

Administration:
- Questionnaire
- No limits on time

Age groups:
- School age children aged 14 – 17: 70 participants
- Adults: 51 participants

Age and prefix turned out to be non-significant
The remaining variables are **acceptability** and **type of word** (standard vs. marginal vs. nonce)

ANOVA RESULTS overall:
F= 546, df = 2, p-value < 2.2e-16

STANDARD WORDS

Words that are standard and conventionalized, might be stored in memory rather than generated on the fly

MARGINAL WORDS

Words that are generated by some speakers and can be understood / accepted by some speakers

NONCE WORDS

Words that cannot be generated and do not exist (because they do not conform phonotactic laws and / or are not based on productive morphological patterns)

ob''jasnit'
uskorit'

oser'ёznit'
uvkusnit'

osurit'
usaglit'

T-test RESULTS for **standard vs. marginal** words:

STANDARD WORDS

Words that are standard and conventionalized, might be stored in memory rather than generated on the fly

MARGINAL WORDS

Words that are generated by some speakers and can be understood / accepted by some speakers

$t = 20$
$df = 21$
p-value = 3.173e-15
95% confidence interval is 277 340

ob''jasnit'
uskorit'

oser'ëznit'
uvkusnit'

T-test RESULTS for **marginal vs. nonce** words:

t = 7
df = 22
p-value =
1.098e-06
95% confidence
interval is 71 135

MARGINAL WORDS

Words that are generated by some speakers and can be understood / accepted by some speakers

NONCE WORDS

Words that cannot be generated and do not exist (because they do not conform phonotactic laws and / or are not based on productive morphological patterns)
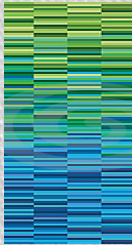
oser'ëznit'
uvkusnit'

osurit'
usaglit'

Hypothesis 3 is confirmed, but...

Hypothesis 1:	Standard	Marginal	Nonce
Hypothesis 2:	Standard	Marginal	Nonce
Hypothesis 3:	Standard	Marginal	Nonce
Hypothesis 4:	Standard	gradual transition	Marginal	gradual transition	Nonce

Marginal words are **much closer to nonce words** than to standard words (compromise between Hypothesis 2 and Hypothesis 3)

Variation across stimuli

Word category	Total score	% of maximal score	Highest score	Lowest score	Interval of variation in scores across stimuli
Standard	11,900	98 %	605	549	56
Marginal	5,729	47 %	479	169	310
Nonce	3,669	30 %	223	150	73

Variation across subjects

Marginal factitive	Gloss	Number of subjects who gave				
		5 points ("normal word")	4 points	3 points	2 points	1 point ("does not exist")
usovremenit'	'modernize'	22	26	27	18	28
opriličit'	'make decent'	9	25	33	22	31

What do these results mean?

- Each type of word has a different behavior
- Marginal words are semantically transparent, but nonce words are not
- **Marginal words** are rated **more like nonce words** than like standard words
- Speakers are **more sensitive to frequency than to semantic transparency**
- **Frequency,** which is related to **performance,** is a **stronger factor than competence** (ability to unpack morphological patterns)
- **Memory** may be a **stronger factor than use of productive rules**