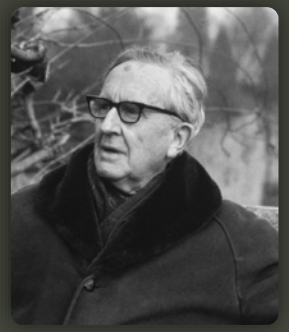
experimental linguopoetics arises

a priori vs *ex post* sound symbolism



could have been Tolkien wrong?

dLCC 2021

jan ‘regis’ havliš

MUNI ARTS

department of linguistics and baltic languages
Brno, Czechia

language creativity in general

linguogenesis (LG) / linguogenerator
: occasional, non-systematic formation of new language phenomena
:: neology, argots, slangs, mixed languages (Czech-German *hantec*)
:: dialects, post-languages (Vulgar Latin > Romance languages)
:: pidgins, creoles

linguopoiesis (LP) / linguopoet
: premeditated, pre-set, systematic formation of new language phenomena
:: e.g. created languages, language purism

products of linguogenesis and puristic linguopoiesis
: initially **not established**, **common** or **widespread**
:: by contact with other potential users get spread & eventually established
:: or dies out

borderland between linguogenesis & linguopoiesis (including pathologic forms)
: neophasia / neophatic polyglossia (compulsory, quasi-systematic formation)
:: idioglossia (so-called private languages)
::: schizophrenian, cryptophasia

experimental linguopoetics in linguistic research

utilisation in cognitive linguistics, psycholinguistics & sociolinguistics

: internalisation of language rules of engineered grammar or language by its acquisition
:: could mind distinguish an engilang from a natlang? *impossible langs* by Andrea Moro
:: will there be a difference when engilang is acquired as L1 or as L2?
:: does mind processes phonetic and morphosyntactic elements differently?

: synthetic / artificial grammar learning (+ serial reaction-time task; AGL-SRT)
:: George A. Miller 1958, Arthur S. Reber 1967

Artificial grammar learning by 1-yr-olds leads to specific & abstract knowledge
Gomez R. L., Gerken L.-A.
Cognition 70, 109-135, 1999

: sentences of constructed grammar – VOT-PEL-JIC, PEL-TAM-PEL-JIC
: test
:: consistent sentences – VOT-PEL-TAM-PEL-JIC
:: inconsistent sentences – VOT-TAM-PEL-RUD-JIC

: synthetic / engineered languages (*engilangs*)

: reversed linguistics
:: construction / creation of a language requires an understanding of language rules
::: language creation is a procedure complementary to its functional description
::: didactic and propaedeutic use to clarify the laws of linguistics

sound symbolism – sound-meaning relation

a priori sound symbolism

strong form
: there is a causal (*a priori*) relation between sound and meaning of a word

weak form
: there is a causal sound-meaning relation within opposite/divergent pairs of meanings
:: small/large, beautiful/ugly, high/low, straight/crooked, etc
:: in many natural languages, there is some sound-meaning relation observed

: **Plato**'s dialogue *Cratylus* (*ca* 4c BCE)
: **Charles de Brosses** (1765)
Traité de la formation méchanique des langues et des principes physiques de l'étymologie
:: strong form proponent – within natlangs, it is their basic principle of formation and development
: **John R. R. Tolkien** (1931) *The Secret Vice*
:: strong form devotee – basic principle of linguopoiesis for his fictional worlds
:: supposedly also within some natural languages (Welsh, Greek or Finnish)
::: e.g. in *Kalevala* epos

Enkä lähe Inkerelle, Penkerelle, pänkerelle
Ihveniä ahvenia, tuimenia, taimenia

ex post sound symbolism

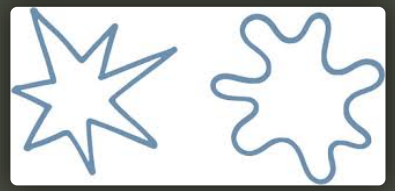
: ***ex post* association of certain sound or sound group to a particular meaning**
:: unimodal & cross-modal imitations, diagrammatic & situational mapping
: it assumes sound-meaning relation reversed to that of the *a priori* sound symbolism
:: certain sounding associates meaning because of a previous sound-meaning connection

: **Otto Jespersen** (1922) *Language – its nature, development and origin*
: **John Rupert Firth** (1930) *Speech*

: *ex post* & *a priori* sound symbolism are considered two possible sound-meaning hypotheses

synaesthesia

cognitively psychological phenomenon
: **experiencing a sensory or cognitive stimulus in a secondary sensory or cognitive pathway**
:: e.g. odours are perceived as colours
:: e.g. shapes perceived as sounds – *kiki*/*bouba* effect



empiric perceptual studies on the sound-meaning relation

study #1 – extensive perceptual *a priori* sound symbolism study

: Czech and Slovak respondents (13)

: choice of a word with assumed meaning based on a sound value (CZ transcription) – 2x 136 choices

:: e.g. which of the two represents meaning **red**: *kizil* – qizil [qɨzil] / *kora* – qora [qɒra]

::: eight meaning pairs (good/evil, snow/rain, red/black, hammer/ needle, bird/animal, fly/fall, ...)

:: each such question is accompanied by a control set **black**: *kizil* / *kora*

:: chosen source languages

::: natural – Hungarian, Irish, Hindi, Zulu

::: *a posteriori* created – Nassian (Slavo-Finnish), Danan (apo-IE)

::: *a priori* created – Arkian, Alurhsa, Itlani, Maltcégj

::: created assumingly following *a priori* s. s. – Quenya, Sindarin, „Cabeian“ (quasi-positive control)

::: randomly generated words – code 1-4 (quasi-negative control)

study #2 – paraphrasing 1994 B. Berlin study – sound symbolism in species nomenclature

: English (39), Czech and Slovak (59), Russian (7) respondents

: choice of the word naming a **fish** based on a sound value (EN, CZ, RU transcriptions)

:: e.g. which of the two is a fish name: pirísh – piríš – пири́ш / kúum – kúum – ку́ум

:: 50 pairs of fish/bird names

study #3 – linguopoetic sound symbolism study

: English (19), Czech and Slovak (21), Russian (5) respondents

: create a word/sound sequence with the given meaning (EN, CZ, RU transcriptions)

:: e.g. create a word in your own imaginary language that you think would represent meaning **red**

::: same meaning pairs as in study #1

study #4 – limited perceptual *a priori* sound symbolism study

: English (17), Czech and Slovak (27), Russian (6) respondents

: choice of a word with assumed meaning based on a sound value (EN, CZ, RU transcriptions)

:: e.g. which of the two represents meaning **red**: *kizil* – qizil [qɨzil] / *kora* – qora [qɒra]

::: eight meaning pairs (good/evil, snow/rain, red/black, hammer/ needle, bird/animal, fly/fall, ...)

:: chosen source languages

::: natural supposedly following *a priori* sound symbolism – Welsh, Greek, Finnish

::: created assumingly following *a priori* sound symbolism – Quenya

::: randomly generated – code 3 (quasi-negative control)

questionnaire data evaluation

blind experiment

: respondents do not know the purpose of the study

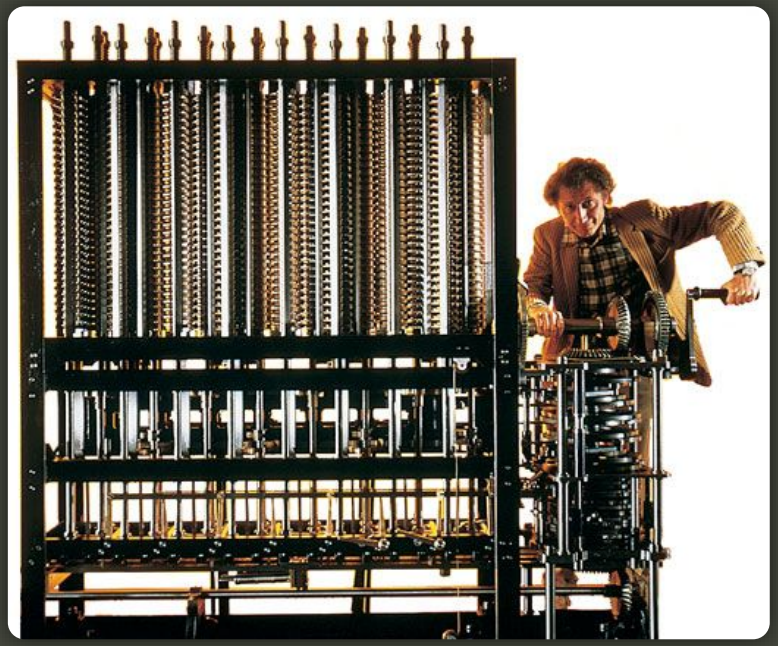
: respondents should not master the languages used

:: mostly works, problems with general education & knowledge

::: Russian, Czech & Slovak respondents recognise Slavic-based langs

::: common knowledge of Greek among scholars ;)

::: some Czech & Russian speakers of Finnish (3)



hypothesis testing

: $H_0 : p_Y = 0.5$, $H_a : p_Y \neq 0.5$

:: simple testing – significancy $0.4 > p_Y > 0.6$)

:: sophisticated testing – z-test, χ^2, McNember test, ...

additional testing

: contingency tables for study #1 (sensitivity, specificity, efficiency)

: 2D result matrix – evaluating probability of choice per word-meaning and per respondent

: phonological analysis of favoured words

results

study #1 – extensive perceptual *a priori* sound symbolism study

: **none of the respondents overall matched significantly meanings with words** ($p_Y > 0.6$)

: 21 out of 136 complementary tasks were significantly correct

:: best results were for Sindarin (6 of 8), the rest was insignificant for a particular language

: 6 of them were significantly incorrect

: respondents (CZ) were successful in recognising words of Slavic and IE origin, but only occasionally

:: *ikala* (NAS) – needle (CZ jehla), *serny* (NAS) – black (CZ černý), ...

:: *neguros* (DAN) – black, *rudros* (DAN) – red, *snygus* (DAN) – snow, ...

: we can find lots of significantly fancy words, which are matched incorrectly (64 out of 272 words)

:: *peto* (DAN), *sulo* (NAS), *osineptu* (CD1), *illas* (CD1), *lutori* (CD1), *dzelášt* (ITL), *ente* (CD2) ($p_Y > 0.8$)

study #2 – sound symbolism in species nomenclature

: **we were not able to get results adequate** to B. Berlin *et al*. 1994 (29 out of 50 significant hits)

:: successful bird names are more of onomatopoetic imitation (*chunchuíkit*, *chichikía*, *takáikit*, ...)

: 4 out of 106 respondents significantly correct at choosing, 1 significantly incorrect

: no significant differences between the three language groups (CZ, RU, EN)

study #3 – linguopoetic sound symbolism study

: **quite a disaster** :-/ instead of ingenious linguopoiesis, 30% of respondents used

:: exotic natural languages : e.g. Japanese, Vietnamese

:: mother tongue biased linguopoiesis : e.g. *bird* – letka (Czech respondent; letět – to fly)

:: L2 (mostly English) biased linguopoiesis : e.g. *to fly* – flájovat, flúga, fúla, fláj, ayra, aérat, volárovat

:: already established conlangs (2)

:: gibberish, mambo jumbo, gobbledegook, jabber, babble, etc. (e.g. oloalao, rarampp, orror)

: the remaining **70%** is **phonologically very heterogeneous** (e.g. *to fly* – fogooryan, vaelah, gah, pellau)

study #4 – limited perceptual *a priori* sound symbolism study

: 5 out of 50 respondents matched overall significantly correctly the meanings to words ($p_Y > 0.6$)

:: 4 out of 50 did it significantly incorrectly ($p_Y < 0.4$)

: **41 respondents were matching overall randomly**

: one respondent excelled in Greek and Finnish (100% matches)

:: ze noted these languages particularly in comment

: 19 out of 27 CZ respondents matched significantly in Quenya, unrivalled by EN (7/17) & RU (3/6)

: respondents matched significantly correctly in Finnish (CZ 15/29; EN 8/17; RU 0/6)

: respondents matched significantly incorrectly in code 3 & Welsh (CZ 18/29; EN 8/17; RU 5/6)

conclusions

we **did not find overall correct choices** between the meanings and the words

we **found a non-random choices** in a small number of meaning-word pairs
: but they were both, correctly and incorrectly matched

: **rejection of null hypothesis** (= there is no word-meaning relation)
:: there are some significantly chosen word-meaning pairs

: **rejection of the *a priori* sound symbolism hypothesis**
:: strong version would require unambiguous overall correct choices
::: there is no such observation
:: weak version would require such correct choices for defined types of meanings
::: there is also no such observation

: **acceptance of the *ex post* sound symbolism hypothesis**
:: there are some attractive words, but not connected to particular meanings
::: but we were not able to find distinctive phonologic pattern (yet)

what remains to be seen...

: **comparing** test results with **spoken** and **written questions** (phonetic notation)
:: a written record of a vocal form of word may be unsuitable for the purposes of the study
::: mind processes vocal and visual aspects of language differently

: **comparing** the effect of **spelling** and **word length** on a choice
:: ťeťem vs tětěm vs tyetyem, želám vs zhelaam; code 4 and Zulu – quite long words

: thorough **phonologic analysis** (Johansson *et al*., *Linguist Typol* 24 (2020) 253)

: thorough **analysis of potential random choices** because of boredom, laziness or lack of interest

aun esse evelienn
de me vorae tháen

dáenace vorae garé
aén tháen!

special thanks to

: John R. R. Tolkien

: Aleš Bičan

: Tony Harris

: James Hopkins

: Benjamin P. Johnson

: Rob Cabe

: Francesco Bravin

: Oksana Shuvalova

: all the respondents