

GENEROVÁNÍ VERŠŮ JAKO PROBLÉM PROZODICKÝ

J I Ř Í L E V Ý — K A R E L P A L A (Brno)

1. Úvodní poznámka

V článku jsou formulovány základní předpoklady a postup práce na některých otázkách české generativní poetiky. V oddíle 2 „Teze J. Levého“ jsou uvedeny výchozí teoretické a metodologické předpoklady. Celý oddíl 2 je uveden v původní a neúplné formulaci bez jakýchkoli změn, smrt zabránila J. Levému provést zamýšlené úpravy. V oddíle 3 „Konceptce“ je charakterizována v hlavních rysech konceptce práce a její jednotlivé části: I, II, III. V oddíle 4 „Současný stav“ se přehledně referuje o první části práce (I), která je už v současnosti realizována. Zatím však nebylo možno provést podrobný rozbor výsledků získaných na samostatném počítači, protože tento úsek práce původně zamýšlel provést J. Levý. Z toho důvodu uvádíme v závěru oddílu 4 jen několik ukázek získaných výsledků. Předpokládáme, že podrobná analýza výsledků získaných v části I bude publikována při vhodné příležitosti později. Autorem oddílů 3 a 4 je K. Pala z katedry českého jazyka na filosofické fakultě UJEP v Brně, který s J. Levým na celé práci spolupracoval (viz vysvětlení v odd. 3).

2. Teze J. Levého

Strukturalistické tendence ve vědních disciplínách zabývajících se složitějšími jevy měly jednu společnou tezi: postavily proti pozitivistické kauzalitě pojem funkce, tj. nesnaží se zjišťovat příčinu jevů, ale jejich pozici ve vyšším celku. Funkční hledisko je nesporně při rozboru systémů účelnější, prostě již proto, že vede ke zkoumání vnitřní stavby systému, a ne jen jejího vztahu k vnějšku (okolí). V první etapě strukturálního zkoumání bylo úskalím funkčního hlediska, že výsledky jím dosažené byly v některých případech nekontrolovatelné, že totiž mnohdy vznikl model, který sice měl svou „vnitřní logiku“, tj. byly správně určeny vztahy mezi jeho prvky, ale nebyla možná kontrola, zda model skutečně vystihuje modelovaný jev. Pozitivisté měli přece jen kontrolu v empirii. Na základě pozorování, že po jevu A následuje vždy jev B, vytvořili genetickou hypotézu, že A je příčinou B, a ta byla verifikována predikcí, tj. očekáváním, že i v budoucnosti po A bude následovat B. Potvrdila-li další pozorování toto očekávání, bylo možno pokládat hypotézu za potvrzenou, bylo učiněno zadost heslu „savoir pour prévoir“.

Funkční hledisko v klasickém strukturalismu postrádalo tuto kontrolu praxí; nebylo často možno ověřit, zda určení funkce prvku v celku skutečně vystihuje

vztahy existující v pozorovaném jevu. Tuto možnost kontroly přináší generativní hledisko; po vzoru generativní gramatiky může také generativní poetika rozložit sdělení do prvků a pravidel pro jejich kombinování takovým způsobem, aby ze zjištěných prvků s použitím generativních pravidel bylo možno text rekonstruovat. Strojový překlad, umělá poezie apod. jsou vedlejším produktem této metodologie, jsou ověřením správnosti zkonstruovaného programu. Tedy pozitivismus vytvářel — a snažil se ověřovat hypotézy genetické, strukturalismus by měl vytvářet a ověřovat hypotézy generativní.

V tomto pojetí je kód deterministický, a proto dokonale formalizovatelný soubor pravidel, jimž musí dílo určitého žánru vyhovovat, aby vůbec šlo např. o český sonet, drama atd. Estetická norma je strategie, z níž vyplývá komplex instrukcí pro výběr jednotlivých stylistických, kompozičních i námětových řešení ze souboru možných alternativ, jež kód připouští. Řečeno naší terminologií: kód je tedy systém definičních instrukcí jednotlivých paradigmat, estetická norma pak soubor instrukcí pro výběr uvnitř paradigmat.

Pro některé typy literatury — zvláště pro verš — dovedeme dnes již kód definovat s dostatečnou přesností, aby např. podle takto definovaných pravidel bylo možno uměle generovat literární text.

Jako příklad jsou uvedena pravidla, jež zapojena do generativní gramatiky češtiny umožňují generovat na samočinném počítači české verše různých metrických forem.

3. K o n c e p c e

Úvodem bychom chtěli upozornit, že následující řádky mají vysvětlující a informativní charakter. Podrobný rozbor zkoumané problematiky bude publikován později, zde se omezíme jen na hrubý nástin a vysvětlíme některé důležité skutečnosti, které vedly k tomuto postupu. Hlavní z nich byla ovšem smrt iniciátora této práce J. Levého († 17. I. 1967). Přesto byla již část výzkumů realizována, o nich bychom chtěli stručně referovat, a dále uvedeme koncepci práce tak, jak byla navržena J. Levým a autorem těchto řádků.

Projekt „Generování veršů jako problém prozodický“ vznikl v r. 1966 během literárně-lingvistických metodologických diskusí mezi J. Levým a autorem z podnětu J. Levého. Projekt má koncepci, která počítá s postupnou realizací jednotlivých na sebe navazujících kroků, přičemž každý z kroků má svůj specifický charakter a může být chápán jako řešení samostatného problému po stránce lingvistické, literární (poetické) i technické (programátorské). Tento postup je metodologicky velmi výhodný, umožňuje pružnou spolupráci a mnohostranné vyhodnocení získaných výsledků. Jednotlivé kroky:

I. Řešení lingvistických a technických problémů, tj. v tomto úseku jde o sestavení nekontextové generativní gramatiky češtiny popisující jazykový systém češtiny (samozřejmě dílčím způsobem; gramatika nepředstavuje vyčerpávající popis jazykového systému) a programu syntézy vět pro samočinný počítač. Program spolu s gramatikou vložený do počítače dovoluje generovat na počítači množinu českých vět, kterou lze v podstatě chápat dvojnásobným způsobem: a) výstup z počítače představuje prozaický text; b) poetický text — formálně volný verš. Bylo by možno postupovat i jinak, tj. zaměřit se přímo na získání výsledků typu b), bez použití generativní gramatiky, jak ostatně někteří

postupují.¹ Soudíme však, že tento postup je metodologicky nekorektní, protože obchází základní stavební kámen všech literárních a poetických útvarů, totiž jazyk, přesněji řečeno popis jazykového systému příslušného přirozeného jazyka. Námí naznačený postup je sice obtížnější (za současného stavu lingvistiky je kromě toho jen částečně realizovatelný), ale podle našeho názoru metodologicky čistý a komplexní (můžeme brát v úvahu problémy syntaktické i sémantické a nezavírají se dveře ani pragmatickému hledisku). Část I byla již realizována, program i gramatika byly v květnu 1967 ověřeny na samočinném počítači SAAB D 21 a byly získány předpokládané výsledky. O struktuře gramatiky a programu se dále stručně zmíníme v oddíle 4.

II. Úpravy programu a jeho doplnění pravidly metrickými a rytmičými, popř. bude-li to nutné, i úprava gramatiky po stránce syntaktické. Autorem těchto pravidel je J. Levý a jsou určena (zatím) pro desetistopý český jamb s rýmy různého druhu. Po sémantické stránce budeme vycházet ze slovníku získaného částečnou excerpcí několika děl Jaroslava Seiferta. Je zřejmé, že část II zahrnuje minimum problémů lingvistických a soustřeďuje se na otázky poetické formy a problémy technické (programátorské). Metrická pravidla jsou budována na vhodné definici české slabiky, která je chápána jako skupina písmen obsahujících samohlásku. Na základě této definice je možno formulovat pravidla strojové identifikace slabiky a pravidla výběru slov o různém počtu slabik ze slovníku stroje, přesněji z gramatiky. Pravidla pro rýmy vycházejí z možností shody slabik, která může být stanovena s předem zadanou přesností, tj. bude možno pracovat s přesně charakterizovaným pojmem kvality rýmu. Kvalita rýmů bude kromě toho úzce souviset s rozsahem slovníku v gramatice (otázka opakování rýmů).

III. Analýza a vyhodnocení výsledků. Jakmile počítač poskytne dostatečné množství poetických textů, bude možno přistoupit k jejich důkladnému rozboru a srovnat je se skutečnými poetickými útvary, jak je naznačeno v tezích J. Levého, které předcházejí našemu výkladu. Je rozumné postavit se na stanovisko, že teze J. Levého reprezentují metodologické předpoklady, z nichž je třeba při analýze výsledků vycházet, ale není vyloučeno, že bude třeba přistoupit k přeformulování těchto předpokladů v závislosti na získaných výsledcích. Zde ostře zasáhla nepsitelná smrt J. Levého; nejsme zatím schopni podrobněji naznačit, k jakým závěrům se dojde v úseku III, jak dalece se budou shodovat nebo lišit myšlenky a záměry Levého se získaným materiálem. Tato konfrontace bude nejobtížnějším úsekem práce, věříme však, že se nám podaří najít vhodného spolupracovníka z řad literárních vědců a celý projekt dokončit. Tím nejlépe uctíme památku J. Levého.

4. Současný stav a stručný přehled problematiky úseku I

Abychom mohli generovat věty přirozeného jazyka (tj. češtiny) na samočinném počítači, musíme mít k dispozici: 1. vhodnou formalizovanou gramatiku, která popisuje (zatím částečně) jazykový systém češtiny; 2. vlastní pro-

¹ Program firmy SAAB zapsaný v jazyce ALGOL-GENIUS, který je schopen generovat švédské a anglické volné verše na počítači SAAB D 21, neobsahuje gramatiku příslušného jazyka, ale pouze jeho slovník. Do současné doby jsme však neměli možnost seznámit se podrobně se strukturou a principem práce tohoto programu.

c e d u r u (algoritmus), která v kombinaci s formalizovanou gramatikou vloženou do počítače produkuje české věty a přiřazuje jim jejich strukturní popisy. V poslední době se k tomuto účelu užívá některých typů tzv. generativních gramatik Chomského, konkrétně gramatik, které se nazývají nekontextové frázové a jsou podrobně popsány v pracích Chomského, Ginsburga aj.² Proto nepokládáme za nutné probírat na tomto místě podrobněji problematiku generativních gramatik a odkazujeme čtenáře na citované práce. Přejdeme přímo k definici nekontextové gramatiky, které je v našem případě použito pro popis úseku českého jazykového systému: $G_1 = (V_T, V_N, R, S)$, kde V_T je množina terminálních symbolů, tj. slovník českých slov, V_N je množina neterminálních symbolů, tj. gramatických kategorií potřebných pro popis češtiny, R je konečná množina přepisovacích pravidel tvaru $A \rightarrow B_1 B_2$, $A \rightarrow C$, $A \rightarrow C_1 \dots C_2$, S je vyznačený výchozí symbol a platí, že $S \in V_N$.

Dále pracujeme s algoritmem syntézy vět, který je v podstatě shodný s algoritmem popsáním u Yngveho.³ Algoritmus využívá principu zásobníkové paměti (pushdown store), takže jedna česká věta je generována během jednoho průchodu algoritmem. Specifikace algoritmu, tj. vlastní program pro počítač, je psán v programovacím jazyce ALGOL-GENIUS. Program byl ověřen a vyzkoušen na samočinném počítači SAAB D 21 v Laboratoři počítačích strojů VUT v Brně.⁴ Ani gramatiku, ani program nebudeme dále podrobněji popisovat, protože jde o čistě lingvistickou a technickou problematiku. Popisy gramatiky a programu jsou publikovány na jiných místech.⁵

Program nyní generuje syntakticky správné české věty, které lze spojit v text, takže výsledkem může být buď prozaický, nebo poetický text (volný verš). Získané věty nejsou vždy z hlediska normální komunikace smysluplné, což může být z estetického hlediska naopak žádoucí. K tomu je třeba poznamenat, že vztah gramatické správnosti a tzv. sémantické smysluplnosti tvoří velký samostatný okruh otázek, jehož prozkoumání může mít z hlediska esteticko-lingvistického velké důsledky. Dosud získané výsledky nebyly však podrobněji prozkoumány a nejsme tedy oprávněni vyslovovat definitivní závěry.

Na závěr uvádíme pro ilustraci ukázky získaných výsledků, jejich podrobný rozbor nebylo však možno do tohoto informativního přehledu zahrnout pro nedostatek času.

Ukázky jsou celkem dvě a velmi jasně charakterizují úroveň dosažených výsledků, které je třeba pokládat za velmi uspokojivé. Ukázka I představuje soubor vět náhodně vybraných ze skupiny 116 vět, které samočinný počítač SAAB D 21 generoval 15. 5. 1967 během 25 minut.

Ukázka II představuje soubor vět vybraných již s jistými tematickými záměry

² N. Chomsky, *Introduction to the Formal Analysis of Natural Languages*, kap. 11, str. 269–321, *Formal Properties of Grammars*, kap. 12, str. 323–418, *Handbook of Mathematical Psychology II*, ed. by Luce, Galanter and Bush, New York 1963; S. R. Ginsburg, *Mathematical theory of Context-free Languages*, New York 1966; S. Kuno, *Computer Analysis of Natural Languages*, Proceedings of the Symposium on Mathematical Aspects of Computer Science, April 5–7, New York 1966 (rotaprint), v tisku.

³ V. H. Yngve, *A Model and an Hypothesis for Language Structure*, Proceedings of the American Philosophical Society, vol. 104, 1960, No 5, str. 444–466.

⁴ Program v jazyce ALGOL-GENIUS zapsala pracovnice Laboratoře počítačích strojů VUT v Brně B. K y l i á n o v á. Používám této příležitosti a děkuji jí za nezištnou pomoc. Současně děkuji vedení Laboratoře VUT za poskytnutí strojového času na samočinném počítači SAAB D 21.

⁵ K. P a l a, *Náhodné generování českých vět*, Slovo a slovesnost 29, 1968, str. 45–56.

ze skupiny 220 vět generovaných 23. 5. 1967 (doba generování 35 min.). V tomto případě byly generované věty řazeny k sobě s cílem vyvolat dojem tematické ucelenosti, takže celá ukázka II má do jisté míry charakter skutečného básnického útvaru. Věty generované počítačem byly do ukázky I a II zařazeny bez jakýchkoli úprav, byla pouze doplněna diakritická znaménka (tiskárna počítače D 21 je vybavena standardní mezinárodní abecedou, která samozřejmě neobsahuje písmena s délkami a háčky), v některých případech byla opravena interpunkce a postavení příklonek.⁶

4.1. U k á z k a I

ZABÍJÍ SE NELÍTOSTNĚ KRÁM
KRADE SE TAM CHLĚB VLAŠSKÉHO ROKU VLAŠSKÉHO
MILENEC EROTICKÉHO STOLU NOVÉHO SNAD NETUŠÍ
BUDOU ŽÁRLIT RYCHLE STEJNÍ HLUPÁCI
ZÍTRA VYTVÁŘÍ CHLĚB, KTERÝ TAM SÁZEJÍ
NEPŘÁTELE SÁZEJÍ RYCHLE
JAK CIZOKRAJNÍ PAVOUCI ŘVALI
BUDE VIDĚT
PRAVÍ PAVOUCI UMÍRALI
A CHLĚB MILOSTNĚ NEMLČÍ
LOVCI PLÁČÍ NEJVÍCE
BUDOU ZABÍJET PRAVÍ CHLAPCI
EROTICKY NEMILUJE
MILOSTNĚ LÉTALI ZEDNÍCI
ANDĚL STOLU ZPITÉHO NEVÍ JIŽ
NEMILUJE
ZADEČEK PROHÁNĚL DŮM PRACHU
BIČ POLODRAHOKAMU EROTICKÉHO PŘIPOMÍNÁ PŘÁVĚ
DŮM SALÓNŮ VAROVAL, ALE HLAS HŘICHU AFRICKÉHO NEPROKLÍNÁ DOBRĚ
ZDĚŠENĚ ČTE MILENEC
BIČ SE VÁBÍ TĚŽKO A NEMILUJE SE
NEJVÍCE PLÁČÍ ZEDNÍCI, MISTR ROZKVETLÉHO ROKU NEMILUJE TAM
BUDOU ZABÍJET SNOUBENCI
PROKŘEHLÍ PRAVÍ MECHANIČTÍ SNOUBENCI BUDOU VYVOLÁVAT
NEJBLIŽŠÍ PAVOUCI CHARAKTER STOLU POPSALI
ZADEČEK ZPITÉHO PŮLNOČNÍHO STOLU TAM CHARAKTERIZUJE
PTÁCI RONILI
STRÍLELI NEJVÍCE NEJBLIŽŠÍ NEPŘÁTELE
MECHANIČTÍ PTÁCI POPSALI DOBRĚ
VYTVÁŘÍ RUČNĚ RYCHLE LÍBEZNÝ CHLĚB
HLUPÁCI ŘVALI
HLUBOKÝ SVĚT STOLU ŘEZALI ČERNÍ NEPŘÁTELE
A BIČ SMYSLU PŮLNOČNÍHO SE ROZNĚCUJE
ZPOZOROVAL ROZVRAT ČAS ZPITÉHO VLAŠSKÉHO ZPITÉHO OSUDU ROZKVETLÉHO

⁶ Pozn. při korektuře (březen 1968): V současné době se připravují experimenty s generováním desetistopého jambu na samočinném počítači, tj. je intenzivně rozpracováván krok II (viz zde str. 75). V době, kdy článek vyjde, budou už pravděpodobně k dispozici první konkrétní výsledky.

BUDOU ODOLÁVAT ČERNÍ NEPŘÁTELE
CO MISTR ZPITÉHO HŘÍCHU NESTAVÍ MILOSTNĚ
VRAZDÍ TAM CHLAPCI
MECHANIČTÍ NEPŘÁTELE TĚŽKO ZPŮSOBÍ
STEJNÍ PAVOUCI BUDOU VYVOLÁVAT NEJVÍCE
JE ZDRČEN ANDĚL, KTERÝ ZPŮSOBÍ
STRÍLELI SE STEJNÍ NEPŘÁTELE
BIČ MILOVAL USTAVIČNĚ
DNES PŮSOBÍ MISTR, KTERÝ ROMANTICKY ROZNĚCUJE
PROBODL CHARAKTER HLAS PRACHU ZPITÉHO
JAK ZMIZELI EPILEPTIČTÍ PTÁCI

4.2. U k á z k a I I

ANDĚL STOLU ZPITÉHO NEVÍ JIŽ
NEMILUJE
EROTICKY NEMILUJE
BIČ POLODRAHOKAMU EROTICKÉHO PRÁVĚ PŘIPOMÍNÁ
A BIČ SMYSLU PŮLNOČNÍHO SE ROZNĚCUJE
BIČ SE VÁBÍ TĚŽKO A NEMILUJE SE
ZDĚŠENĚ ČTE MILENEC
MILENEC EROTICKÉHO STOLU NOVÉHO SNAD NETUŠÍ
BUDOU ZABÍJET PRAVÍ CHLAPCI
VRAZDÍ SNAD EPILEPTIČTÍ MECHANIČTÍ PTÁCI
VRAZDÍ TAM CHLAPCI
BUDOU ZABÍJET PRAVÍ CHLAPCI
LOVCI PLÁČÍ NEJVÍCE
PRÁVĚ PLÁČÍ MECHANIČTÍ STEJNÍ NEPŘÁTELE
STRÍLELI NEJVÍCE NEJBLIŽŠÍ NEPŘÁTELE
VAROVAL HLAS JEDNOHO SMYSLU NOVÉHO
JE ZDRČEN ANDĚL, KTERÝ ZPŮSOBÍ
ČAS EROTICKÉHO ROZKVELTÉHO ROZKVELLÉHO VLAŠSKÉHO POLODRAHOKAMU
SE
MALUJE MARNĚ
PLYNE MARNĚ CHLĚB
BIČ SE VÁBÍ TĚŽKO
NEMILUJE SE
PROKŘEHLÍ HLUPÁCI MILOSTNĚ TANCUJÍ
CO ROMANTICKY VIDÍ RŮŽOVÝ MISTR
VZTYČIL BIČ ROZKVELTÉHO EROTICKÉHO ROZMARU ROZKVELTÉHO
A BIČ SMYSLU PŮLNOČNÍHO SE ROZNĚCUJE
NELÍTOSTNĚ NEPROKLÍNÁ
ANDĚL, KTERÝ VRAZDÍ OBLÁČEK ROZMARU, ROZNĚCUJE JEDNÍM ZPŮSOBEM
ZVLÁŠTNÍM
PŘIPOMÍNÁ
HLAS, KTERÝ HLEDALI USTAVIČNĚ, ZDĚŠENĚ NEMLČÍ
JE OVLÁDÁN HLAS AFRICKÉHO OSUDU
HLAS, KTERÝ KUPOVÁVAL DŘÍVE,

POZNAL
CIZOKRAJNÍ STEJNÍ LOVCI UMÍRAJÍ
VRAŽDÍ
NEBUDE ODOLÁVAT ÚZKÝ DŮM
DŮM SALÓNU VAROVAL
ALE HLAS HŘÍCHU AFRICKÉHO NEPROKLÍNÁ DOBŘE
MILOVAL LEHKÝ SVĚT VZORU PYŠNÝ ANDĚL
BIČ EROTICKÉHO VLAŠSKÉHO PŮLNOČNÍHO HŘÍCHU NOVÉHO VZTYČIL NEJVÍCE
TANCUJÍ ROZVRAT HŘÍCHU EPILEPTIČTÍ CHLAPCI
JE ZDRČEN ANDĚL, KTERÝ ZPŮSOBÍ
ČAS SE ZPOZOROVAL SNAD
BUDE VIDĚT MILENEC NOVÉHO POLODRAHOKAMU
ANDĚL HŘÍCHU NOVÉHO ASI OBRACÍ
PŘIPOMÍNÁ OSOBITÝM ZPŮSOBEM SVATÝM MILENEC
JE ZAHALEN ČAS VLAŠSKÉHO HŘÍCHU VLAŠSKÉHO
BUDE VIDĚT
JAK MILENEC AFRICKÉHO ZPITÉHO PŮLNOČNÍHO ROKU ROZKVETLÉHO SE
PŘIPOMÍNÁ
MILOSTNĚ
ASI OBRACÍ ŠŤASTNÝ ČAS
NEVÍ SE

ПОРОЖДЕНИЕ СТИХОВ КАК ПРОСОДИЧЕСКАЯ ПРОБЛЕМА

Настоящая статья посвящена основным предпосылкам и методологическим вопросам, возникающим в связи с разработкой некоторых проблем чешской порождающей поэтики.

Теоретическое и методологическое ядро работы содержится в части 2, автором которой является Й. Левы и которую он уже не успел закончить (Й. Левы умер 17. 1. 1967).

Часть 3 содержит характеристику концепции работы и описание отдельных шагов, последовательность которых представляет решение всей проблемы. Шаг I включает решение лингвистической проблематики, т. е. составление неполной контекстно-свободной порождающей грамматики, описывающей систему чешского языка, и составление программы для синтеза чешских предложений на электронной вычислительной машине. Программа вместе с грамматикой позволяет вычислительной машине порождать множество чешских предложений, которое можно понимать: а) как прозаический текст; б) как поэтический текст имеющий форму вольного стиха. Шаг I в настоящее время уже осуществлен, программа вместе с грамматикой удовлетворительно работают и были проверены на вычислительной машине СААБ Д 21 в течение мая 1967 г. Результаты опытов (I и II) приведены в конце статьи (с. 46—47). Программа для вычислительной машины записана на программирующем языке АЛГОЛ-ГЕНИУС.

Шаг II содержит разработку метрических и ритмических правил и включение их в программу для вычислительной машины. Эти правила предложил Й. Левы, и мы предполагаем, что на их основе будет можно получить при помощи вычислительной машины 10-стопный чешский ямб с рифмами разного рода. Ритмические правила основаны на определенной дефиниции чешского слога, понимаемого как группа букв содержащая гласную.

В шаге III мы собираемся перейти к подробному анализу результатов полученных вычислительной машины и к их сравнению с реальными поэтическими текстами.

Заметим, что статью в настоящем виде следует считать первой и наглядной информацией о текущей работе на проблемах чешской порождающей поэтики. Подробный анализ пока полученных результатов будет опубликован позже.

Статья была написана уже после смерти Й. Левого, который является автором незаконченной части 2, автором остальных частей (1, 3, 4) является К. Пала из кафедры чешского языка Философского факультета в Брно, который сотрудничал с Й. Левым на указанной проблематике.