

Gépi fordítási munkálatok Bulgáriában

Bármely tudományos eszme keletkezése, gyakorlati megvalósítása a szükséges anyagi és elméleti előfeltételek fokozatos létrejöttének eredménye. Így alakult ki világszerte az ötvenes években a gépi fordítás. Az a körülmény, hogy a gépi fordítási munkálatok több országban, így Bulgáriában is, akkor indultak meg, amikor a gépi fordítás elmélete és gyakorlata jelentőset fejlődött, megkönnyítette több feladat megoldását, bizonyos fázisokon gyorsabban túljuthattunk és fokozatosan áttérhettünk a világszínvonalon folyó munkára.

A bulgáriai gépi fordítás története rövid és még csak a kezdeteknél tart, fejlődésében néhány szakaszt állapíthatunk meg, amelyek rövid vázolósa cikkünk feladata.

I. A gépi fordítás gondolatának létrejöttét és gyakorlati megvalósulását korunk tudományos és technikai eredményei, elsősorban a nyelvelmélet, kibernetika és a fordításelmélet fejlődése tette lehetővé.

DE SAUSSURE eszméi Bulgáriában elsősorban IVÁN LEKOV professzor munkásságára voltak hatással, aki a funkcionális strukturalista prágai fonológiai iskola híve volt.¹ IVÁN LEKOV első munkáiban megállapítja, hogy szükség van az újgrammatikus iskola tanainak revíziójára, így például a hangtörvények jellegének megítélésében; hangsúlyozza a kontamináció szerepét stb.² A fonológia kérdésköréből³ foglalkozik a fonológiai oppozíciók stabilitásával és funkciójával, statisztikai módszereket alkalmaz a kutatásaiban és a nyelvi tényeket diachronikus aspektusban elemzi.⁴

A tipológiai kutatások megteremtették a nyelvi tények kvantitatív vizsgálatának előfeltételeit,⁵ és lehetővé vált, hogy nyelvtudásaink fokozatosan azokhoz a modern irányzatokhoz csatlakozzanak, amelyek a nyelvtudományban egzakt módszereket alkalmaznak. E téren is I. LEKOV tette meg az első lépéseket, ennek eredményeképpen lehetővé vált a nyelvészek, kiberneti-

¹ *Iv. Lekov*, Нови насоки в съвременното езикознание, Философски преглед, IX (1937), кн. 2.

² *Iv. Lekov*, Прояви на фонологична свръхстарателност в развоја на българския, полския и чешкия език. Списание на българската академия на науките, кн. XLIII. 1939.

³ *Iv. Lekov*, Фонологичната стойност на удължените и удвоени съгласни звукове в славянските езици. Годишник на С У, Историко-филологически факултет, XXXVI (1940), кн. 4. 1—110.

⁴ *Iv. Lekov*, За диалектиката в езиковата система (биномният срещуположен характер на граматическата структура по данните на славянските езици). Годишник на С У, Историко-филологически факултет, XLII (1947—1948), 1—33.

⁵ *Iv. Lekov*, За общия характер на количествените промени в българския език. Български език, X (1960), кн. 4. 289—296.

kusok gyümölcsöző együttműködése a nyelvnek mint szemiótikai rendszernek a tanulmányozása terén, elterjednek a matematikai statisztika, a valószínűség-számítás, halmazelmélet módszerei.

Ugyanakkor a bolgár nyelv tudósok egyre határozottabban törekсенek a nyelvészeti alapfogalmak meghatározására, a nyelvstatisztikai módszer segítségével a nyelvi tények következetes vizsgálatára. Meg kell említenünk M. JANAKIEV docens cikkét „Azokról a grammémákról, amelyeket a bolgár nyelv tanban »jelen időnek«, »jövő időnek« szokás nevezni”,⁶ és „Nyelv, kód, stílus, nyelvjárás” című tanulmányát,⁷ amelyben kód fogalmát természetes nyelvekre alkalmazza. C. MLADENOV a bolgár fonetika néhány fogalmát igyekszik meghatározni a matematikai logika terminusaival.⁸ Ezzel párhuzamosan megjelennek olyan cikkek is bolgár szerzők tollából, amelyek strukturális analízissel, generatív grammatikákkal és transzformációs elemzéssel foglalkoznak.⁹

1963-ra a bolgár matematikusok és technikusok elkészítik a B 1 számítógépet. Így létrejön a további kibernetikai kutatások technikai bázisa, bővülnek a Számítástechnikai Központ feladatai, a Matematikai statisztikai és valószínűség-számítási osztály célul tűzi ki az írott nyelv statisztikai paramétereinek, a nyomtatott szöveg entrópiájának vizsgálatát.¹⁰

A fordítás, elsősorban a műfordítás kérdéseivel a bolgár filológusok már elég régóta foglalkoznak. Az ötvenes évek elején a fordítási munkálatok rohamos fejlődése során számos hiányosságra figyelnek föl, és felmerül annak szükségessége, hogy a problémákat elméleti úton kell megoldani.¹¹ A LJUDSZKANOV cikke „A funkcionális ekvivalensek elve a fordításelmélet alapja” a fordítási folyamatot elemezve megállapítja különböző műfajok azonos vonásait, leszögezi egy külön nyelvészeti diszciplína, a fordításelmélet szükségességét és lehetőségét. Véleménye szerint bármely szakfordítás egységes módon írható le, s a fordításelmélet feladata, hogy megállapítsa a különböző szövegek fordításának közös törvényszerűségeit, és — a nyelvi közlés fogalmából — a céljából kiindulva — meg kell határozni a fordításelmélet tárgyát.¹²

Az elmondottak eredményeként a hatvanas évek elejére létrejönnek a gépi fordítási munkálatok feltételei Bulgáriában. Elsőrendű feladatunknak tekintettük, hogy a bolgár tudományos közvéleményt megismertessük a gépi fordítással, bebizonyítsuk szükségességét, és a gépi fordítás elméleti alapjaként megalkossuk a fordítás általános tudományos elméletét.

⁶ М. Янакиев, За грамемите, наричани в българската граматика «сегашно време» и «бъдеще време». Сборник в чест на проф. К. Мирчев. Известия на института за български език, кн. VIII, С., 1962.

⁷ М. Янакиев, Език, код, стил, диалект. ЕЛ, 1961/4.

⁸ Ц. Младенов, Бележки върху римовото движение в стиховата реч: Славянска филология, т. 4, 1963; Уб. Един генериращ модел на римовото движение в стиховата реч, ЕЛ, 1963/6. Съвременната българска азбука като семиотична система, ЕЛ, 1964/4.

⁹ Ц. Младенов, Традиционният синтактичен анализ като аксиоматична система. ЕЛ, 1964/5.

¹⁰ В. PENKOV, А. OBRETOV, В. GENDOV, Т. KIRPINOVA and Т. JOULEANOV, Frequencies of Letters in Written Bulgarian. Доклады Болгарской Академии Наук, 15, № 3, 1962.

¹¹ А. Людсканов, Принципът на функционалните еквиваленти — основа на теорията на превода. ЕЛ, 1958/5.

¹² А. Людсканов, Введение в теорию перевода А. В. Федорова. (ismertetés) ЕЛ, 1958/6; За степеня на точността на превода. ЕЛ, 1959/1.

2. 1960 nyaratól kezdve jelenik meg A. LJUDSKANOV cikksorozata „Az algoritmus fogalma és jelentősége a gépi fordítás számára” címmel.¹³ Bennük a szerző vázolja a gépi fordítás történetét, elemzi a fordítás folyamatát, bevezeti az invariáns információ fogalmát, ami minden fordítás célja, levezeti az önálló gépi fordítás meghatározását, és megvilágítja, miért van szükség gépi fordításra Bulgáriában. Ebben az évben jelenik meg I. LEKOV cikke a modern nyelvtudományi irányzatokról,¹⁴ amely hangsúlyozza az alkalmazott nyelvészet, elsősorban a gépi fordítás fontosságát, amely egyszersmind az elméleti nyelvtudomány kísérleti bázisa is. Ettől kezdve az „Ezük i literatura” folyóirat rendszeresen közöl modern nyelvészeti cikkeket, amivel jelentősen elősegíti a gépi fordítás fejlődését.

E sorok írója két fordításelméleti cikkében¹⁵ a kibernetikai szemléletre alapozva, információelméleti és szemiótikai koncepciók segítségével az emberi és gépi fordítás fajtáit egy közös fogalomból származtatja: $L_1^M \rightarrow L_2^M$. Az $L_1^M \rightarrow L_2^M$ a belső szemiótikai fordítás ($L_1^x \rightarrow L_2^x$) egy faja, amely a szemiótikai rendszerek közötti fordítással együtt ($L_1^x \rightarrow L_2^y$) egy közös fogalomhoz: $L_1 \rightarrow L_2$ tartozik, ami bármely kódról bármely más kódra való áttérést szimbolizál.

A gépi fordítás eszméjének népszerűsítése, a fordítás tudományos koncepciójának kialakulása, a nyelvelmélet és gépi fordítás kapcsolata lehetővé tette a tulajdonképpeni gépi fordítási munkálatok megindulását.

3. A bolgár gépi fordítási munkálatokra ösztönzően hatott az a körülmény, hogy a szófiai nemzetközi szlavista kongresszus tematikájában — ezúttal először — helyet kapott az alkalmazott nyelvészet. Így került sor az orosz — bolgár gépi fordítás néhány kérdésének kidolgozására. Az első egy olyan általános sémát tartalmaz, amely megoldja az orosz nyelv szintetikus és a bolgár analitikus szerkezetéből adódó problémákat.¹⁶ A másik az orosz — bolgár gépi fordítási szakszótárak összeállításának elveit állapítja meg.¹⁷

Az orosz — bolgár gépi fordítás konkrét kérdéseinek vizsgálata során kiderült, hogy szükség volna egy olyan munkára, amely tisztázza a bolgár gépi fordítás elveit, és megszabja fejlesztésének útját. Ezt még aláhúzta az a tény, hogy nincs a nemzetközi irodalomban olyan munka, amely a gépi fordítás elméletét összegezné (nem számítva az elavultakat). Ezt a hiányt volt hivatva pótolni e sorok írójának monográfiája „A gépi fordítás elméletének alapjai és az orosz — bolgár gépi fordítás”.¹⁸ A munka kritikai áttekintést nyújt a legfontosabb gépi fordítási irányzatokról, bemutatja az orosz — bolgár gépi fordításhoz szükséges szótári információk paramétereit, a szótárak,

¹³ A. Людсканов, Понятието алгоритъм и неговото значение при машинния превод. ЕЛ, 1960/2, 4, 6 és 1961/1; lásd még: Пренебрегваме ли кибернетиката? Народна култура, № 4, 1, 1962; Кибернетиката и задачите на университета. Софийски университет, 1962/4.

¹⁴ Ив. Левков, За някои съвременни черти на езиковедската теория и практика, Език и литература, 1960/4. (A továbbiakban ЕЛ.)

¹⁵ A. Людсканов, Към въпроса за предмета на общата теория на превода. Год. фф на СУ, 1963; За предмета, мястото, систематиката и методологията на общата теория на превода. т. I/414, т. II/194.

¹⁶ A. Людсканов, Отговори на въпросите от научната анкета по езикознание. т. 3, С. 1963.

¹⁷ A. Людсканов, Някои бележки относно общите принципи за съставяне на речник за руско-български МП. Славистични студии, С. 1963.

¹⁸ A. Людсканов, Основи на теорията на машинния превод с орлед на руско-българския МП. Год. фф на СУ, 1964, 220.

a lexikai, alaktani és mondattani elemző algoritmus szabványosított formáit. Elemzi a többértelműség vizsgálatához szükséges sémát, amely minden szinten használható,¹⁹ és rámutat arra, hogy be kell vezetni a „szükséges fordítási információ”: I(T) fogalmát. Az egyetemes gépi fordítást megelőző lépésként lehetőség van arra, hogy bináris (ez esetben orosz—bolgár) fordításban megállapíthassuk azt a szükséges és elegendő információt, amelynek segítségével a bemenő közlés nyelvi komponenseinek jelentése egyöntetűen meghatározható és gráfok segítségével ábrázolható. Így a gépi elemzés során az adott komponenshez megadjuk az összes elvileg lehetséges és szükséges paramétert, és később töröljük a szükségteleneket. (Ehhez hasonló Lecerfe „konfliktusos módszer”.) A munka végén a szerző vázolja a bolgár gépi fordítási munkálatok programját, hangsúlyozza az értelmi elemzés jelentőségét.²⁰

Ezzel egy időben indul meg a szófiai egyetemen a „Gépi fordítás alapjai” szakkolégium, és az orosz—bolgár gépi fordítás egyes részletkérdéseit szakdolgozatként adják ki.

4. 1964 őszén A. LJUDSZKANOV vezetésével megalakult a Gépi fordítási Csoport a Bolgár Tudományos Akadémia Számítástechnikai Központjának és Matematikai Intézetének a Matematikai Statisztikai osztályán. Az év végéig megvitatásra kerültek a csoport távlati tervei, és megkezdődött a kísérleti munka. A következő alapfeladatok megoldására van szükség: matematikai szövegek szótárának összeállítása (ami eddig kb. 2500 szót tartalmaz, s további bővülése során a fordító rész függetlenné válik a bemenőtől); a szótári információ paramétereinek megállapítása; a leírás szabványosítása (elkészült); a lexikai elemzés általános formájának kialakítása; az orosz tudományos és technikai szövegek szó szerinti fordításának kidolgozása; a szó- és morféma-szerinti fordítás lehetőségeinek megvizsgálása; a független bolgár szintézis első lépéseként elkészítendő a szórend és az artikulus sémái; elvégzendő politikai szöveg kísérleti fordítása (ehhez egy 1000 szavas szótár, alaktani és strukturális mondattani elemző algoritmus készül, amely később teljesen strukturális mondattani, majd szemantikai alapra kerül, hogy megállapíthatók legyenek az egyes módszerek előnyei). Mindezek lehetővé teszik a bolgártól független orosz elemzés, a független bolgár szintézis optimális rendszerének kialakítását, s idővel a más nyelvekről történő fordítást.

Ezzel párhuzamosan kísérleti munka folyik a Minszk-2 gépen, lefordították egy orosz tudományos szöveg néhány mondatát, és előkészítjük hasonló francia szöveg gépi fordítását is. Aránylag nagy anyagon kipróbáltuk a lexikai elemzés egyik lépését, a bemenő szöveg elemzését mondatokra, aminek eredménye a mondatok elhelyezése az operatív memóriában. Ehhez elkészítettük a rövidítések kis szótárát. A második séma, a mondat felosztása szavakra, ezután lép működésbe; ezt két változatban dolgoztuk ki, az elsőben a bemenő szöveg perforálásánál közel jeleztük azokat az írásjeleket, amelyeket nem választ el köz a környező szövegtől, a másodiknál nem alkalmaztunk jelzést. A kielemezett szavak bekerültek az operatív memóriába. Mindkét séma 106 utasításból áll. A harmadik séma „Tulajdonnevek” a szó-

¹⁹ A. Людсканов, Към въпроса за обща схема «Многозначност» при машинния превод: Сборник в чест на 60 г. на проф. И. Лекон. Известия на Института за български език, 1964/XI.

²⁰ A. Людсканов, Смеслов подход при машинния превод. ЕЛ, 1964/2.

tári keresés után lép működésbe. Ha ugyanis van „meg nem talált” elem, a gép formális kritériumok alapján megállapítja, hogy ez formula, meg nem talált szó vagy tulajdonnév. Az első esetben kigépelem, a másodikban a mondatban megmarad az orosz szó száma, és a gép megjegyzi ezt a szótár további kiegészítése végett. A harmadik esetben a gép elkezd a tulajdonnév fel dolgozását: az egy vagy több betűs rövidített tulajdonneveket kigépelem a mondat megfelelő helyére, a nem rövidítetteket alakitanilag elemzi. S_g -operációval, tehát jobbról balra haladva, levágja a végződést, összeveti az alakítani szótár adataival és bolgár megfelelőjét hozzáteszi a tőhöz. Ezenkívül szakdolgozatként kidolgozzák a következő sémákat: „A főnévi tő mássalhangzóváltakozásai”, „Állandó szókapcsolatok”, „Összetett szavak” (orosz matematikai szövegekben).

A végzett kísérletek lehetővé tették két elvi kérdés eldöntését. Először: elméleti és gyakorlati követelményeket összevetve milyen mértékben alakításuk perforálásnál a szöveget (mondatokra, szavakra bontásnál, írásjelek, formulák, idegen idézetek stb. esetén). Másodsor: szabvány programokat kell összeállítani a gépi fordításhoz, s a tudományágak szótárai mellett össze kell gyűjteni a közös szókincset is.

5. A tulajdonképpeni gépi fordításon kívül több vele kapcsolatos nyelvészeti, statisztikai munkálat folyik.

A BTA Bolgár Nyelvészeti Intézetének szlavisztikai osztályán a szóhosszúságot elemzik, ezenkívül az osztály vezetője, IVÁN LEKOV megvizsgálta a szláv nyelvek fonémáinak, morfémaiknak és szavainak kapcsolódását,²¹ s a mondattani struktúra kérdéseit. Szerinte alapvető szintaktikai egység a szintakszéma, amely összevethető a többi szint alapegységeivel. R. MUTAFCSIEV, az Intézet tudományos főmunkatársa a bolgár szövegek szófaji eloszlását vizsgálja a stíluselemzés szemszögéből.²²

A BTA Matematikai Intézetében 1962-ben megindult nyelvészeti szeminárium folytatja munkáját, itt olyan előadások hangzanak el, amelyek mind a nyelvészek, mind a matematikusok érdeklődésére számot tartanak.

M. JANAKIEV vezetésével elkezdődtek a bolgár gyakorisági szótár munkálatai is. Elkészült a bemenő szöveg szegmentálására szolgáló formális utasítás, amelynek segítségével megállapíthatók a fonetikus szavak, és megadhatók bizonyos kiegészítő információk (mondathangsúly, pauza, fonemikus átírás, amely a fonéma helyzeti variánsait is figyelembe veszi). Az így kiválasztott fonetikai szók és kiegészítő információk kerülnek a lyukkártyára. Az osztályozás eredményeképpen megkapjuk a gyakorisági szótár szójegyzékét, amely feltünteti a leggyakoribb fonetikai szavak gyakoriságát és a különböző hosszúságú szók statisztikai eloszlását. A szófiai egyetem filológiai karán több szakdolgozat készül, amely a gyakorisági szótár munkálatait segíti elő. 1965 őszére körülbelül 3 millió grafikai szót tartalmazó szöveg szógyakorisági viszonyairól kapunk képet.

²¹ *Ив. Лекков, Към въпроса за съчетаемостта (валентността) на частите на речта в българското изречение. Отázкэ slovanské syntaxe, Sbornik Brnenské syntaktické konference. Praha, 1962. 47—57.*

²² *Р. Мутафчиев, Няколко статистически наблюдения върху стила на мемоара в последното десетилетие на миналия век: Славистически сборник. С. 1963; Уб., Статистическо распределение на глаголни форми ЕЛ, 1965/2.*

A Külföldi Hallgatók Intézetében (Szófia) megtörténtek az első lépések a programozó nyelvtanulás kidolgozására.

A gépi fordítás kérdéseit az 1965. évtől felvették a Műszaki és Tudományos Fejlesztési Bizottság tudományos terveibe.

*

A továbbiakban jobban össze kívánjuk hangolni a különböző helyeken folyó gépi fordítással kapcsolatos munkálatokat, amelyek hozzájárulnak marxista nyelvtudományunk további fejlődéséhez.

A. LJUDSZKANOV