

rek olyan fordításokat szolgáltatnak, melyekhez a szükséges utószerkesztési és lektorálási munka elfogadható keretek között marad és a nyersfordítás minősége is elfogadható lesz. A lektor munkája pedig annyiban változik, hogy azt ún. *interaktív rendszerek, azaz ember-gép dialógusok segítik és gyorsítják.*

Az, hogy a minőségi fordításoknál az emberi munka teljesen kiküszöbölhető legyen, már csak a fordítandó szövegek sokrétűsége és eltérő minősége miatt sem lehetséges. Elképzelhető, hogy meghatározott célokra egységes előírások szerint készült, kötött szövegek kerülnek fordításra. A textilipari referátumok fordításánál ez a módszer már jól bevált. Kétséges azonban, hogy hasonló megkötések tudományos és műszaki szövegek esetében keresztülvihetők lehetnek. Erre irányuló amerikai kísérletek már korábban kudarcba fulladtak.

A képezett fordítók szerepe

Ha feltételezzük, hogy a nyersfordítások a jövőben gépi úton fognak készülni, akkor a nagyobb követelményeket támaztó lektori munkára mind a fordítóknak, mind a fordítókat képző intézményeknek idejében fel kell készülni. Nem valószínű, hogy a számítógép elveszi a fordítók kenyerét, hiszen a fordítóirodák általános tapasztalata az, hogy kevés a valóban kiváló nyelv- és szaktudású fordító. *Kívánatos tehát a fordítók, ill. az utánpótlás képzési szintjének emelése.*

A számítógép nyilvánvalóan kivívja helyét a fordítás területén, akár gépi fordításról, akár gép közreműködésével készült fordításról beszélünk majd. Hogy mennyire lesz lehetséges a minőségileg megfelelő fordítások átfutási idejének lerövidítése, az nagymértékben a lektoroktól függ. Az óránkénti 300 ezer szó – ami a gépi fordítási rendszerek teljesítménye – mintegy 850 A/4 oldalt jelent. Ha – jelenlegi dfázisukat is figyelembe véve – a lektorok óránként 1–3 oldalt néznek, illetve szerkesztenek át, akkor a gépi és az emberi teljesítmény aránya 1 : 850, ill. legjobb esetben 1 : 280.

Fordítandó szövegekben nincs hiány, a gép el lenne látva munkával. Kétséges azonban, hogy a képezett fordítóknak mutatózó és ezt kiegészítő szükséglet egyáltalában fedezhető-e. A gépi fordítások mennyiségi növelésének tehát nem sok értelme van, inkább *a minőség javítására kellene törekedni* – ehhez azonban még hosszú és nehéz út vezet.

/BRINKMANN, K. H.: Perspektiven der maschinellen Übersetzung. = Nachrichten für Dokumentation, 29. köt. 3. sz. 1978. p. 104–108./

(Dezső Zsigmondné)



Többsnyelvűség a tudományos és műszaki információban

Nyelvi problémák és következményeik az információs hálózatban

A dokumentációs rendszerek gépesítésének következményeképpen *az adatbázisok egymással szoros kapcsolatba fonódó, nemzetközi hálózatokká válnak.* Ezáltal azonban a hálózat összetétele meglehetősen heterogén lesz, és nehezen küszöbölhetők ki a párhuzamosságok. A tudomány és technika fejlődése következtében a mind több új nyelvterületen kialakuló információs központok megdöntik azt a néhány év előtti elképzelést, miszerint egyetlen nyelv átveheti az általános kommunikációs eszköz szerepét.

A tudósok már nem egyetlen fejlett országban élnek, hanem elszórtan dolgoznak a világ legkülönbözőbb nyelvterületein. Ha a szükséges információk csupán egyetlen, számukra idegen nyelven állnak rendelkezésre, e nyelv elsajátítása szerves részét képezi a szakmai munkához tartozó tájékozódásnak, bár sok időt és energiát vesz igénybe. Lényegesen könnyebb a helyzet, ha a tudóst jól szervezett hazai információs szolgálat saját anyanyelvén segíti munkájában. Az ehhez szükséges hagyományos fordítói munka azonban nehézkes és költséges.

Gépi fordítási kísérletek

Tekintettel a Szovjetunió bő és értékes szakirodalmára és arra a tényre, hogy a nyugati országokban kevés tudós bírja az orosz nyelvet, az Egyesült Államokban, Európában és Kanadában számos kísérletet folytattak az oroszról angolra (és később franciára) történő gépi fordítás megoldására. Viszonylag egyszerű információkkal, mint pl. meteorológiai jelentésekkel már kitűnő eredményt értek el. Hasonlóan egyszerű információknak tekinthetők a referátumok is. A textilipar területén TITUS I., II. ill. III. néven már működnek különböző fokozatai egy ilyen jellegű, többsnyelvű dokumentációs rendszernek. Emellett Franciaországban, Kanadában és az NSZK-ban is kidolgoztak már olyan számítógépi programokat, amelyek kielégítő eredménnyel végzik a tudomány és technika területén készült referátumok gépi fordítását.

Az információs gépi fordítás jövője

Az eredmények ellenére a gépi fordítás még nem tökéletes; hibák, félreértések még mindig előfordulnak. Ezeknek a javítás érdekében történő elemzése folyik.

A leggyakoribb hibák**a) Helytelen mondat szerkezet, logikai félreértések**

Ez a hiba a szavak összekapcsolásánál, a kötőszavak (pl. a „hogymint”) megfelelő helyének megválasztásánál fordul elő. A megoldás olyan kompromisszum, amely a csupán legkönnyebben feloldható félreértéseket zárja ki, esetleg kétes esetekben a gép kinyomtatja a különböző változatokat.

b) Nyelvtani, szófaji bizonytalanság

Nemcsak a homonimák okozzák, hanem az azonos alakú szavak szófaji eltérései is, amikor ugyanaz a szó éppen úgy lehet ige, mint főnév vagy melléknév. Az ilyen jellegű hibák kizárása azonban viszonylag könnyebb.

c) Nyelvi fordulatok, sajátos kifejezési formák

Az ún. idiómáknak az értelme vezethető le a bennük foglalt egyes szavak értelméből, szó szerinti fordításuk tehát értelmetlen, teljességgel félreérthető. Számuk viszonylag kicsi, de gyakran fordulnak elő. Egyetlen megoldás jegyzékbe foglalásuk és előzetes lefordításuk.

A különféle hibák előfordulási gyakorisága

A nehézségekről némi képet ad az a megállapítás, mely szerint a felmerülő lehetséges hibák 20%-a a fenti három kategóriába tartozik. Ebből 12% a helytelen mondat szerkezet, 4% a nyelvtani, szófaji bizonytalanság, 4% pedig a nyelvi fordulatok kategóriáját érinti. Ezekből csupán az a) kategóriába tartozó hibák kiküszöbölése jár nagyobb költséggel. Második generációjú analízátor segítségével csaknem valamennyi b) és c) kategóriájú hiba és az a) kategóriájú hibák nagy része jelentősebb kiadások nélkül is elkerülhető.

Néhány jelentéktelenebb hiba

a) *Főnév vagy szó szerkezet*, melyre személyes vagy vonatkozó névmás vonatkozik.

Az európai nyelvekben általában elég világosan kitűnik, kire vagy mire vonatkozik a névmás, de előfordul, hogy a fordításban „ő” vagy „az” egyaránt szerepel. Félreértést okozhat a meg nem felelő szórend is, amikor a főnév és a rá vonatkozó névmás távolsága – illetve egy másik főnév közbeiktatása – megváltoztatja a vonatkozás értelmét. Szerencsére az efféle hibalehetőség a referátumokban ritkán (az össz hibák 1%-ában) adódik. Amennyiben mégis előfordul, a hibás fordítás következménye olyan képtelenség, mely már eleve kizár minden félreértést.

b) Többértelműség

Akkor fordul elő, amikor valamely főnévnek (néha igeinek) a célnyelvben több megfelelője is van. A félreértés kiküszöbölése főleg a dokumentátorra hárul, de feltétlenül szükség van a rendszer rugalmasságára is, amennyiben menetközben kell a szótárból kiemelni,

illetve abban más szavakkal helyettesíteni a félreértésekre alkalmas adó kifejezéseket.

c) Prepozíciók és névelők

Tekintve ezek összetett és változó alkalmazását az európai nyelvekben, az ilyen jellegű hibák okozzák a legtöbb nehézséget. A tudományos információban a jelentésárnyalatok általában kerülendők, így az információt kevésbé torzítják effajta hibák. Ennek ellenére mindent el kell követni kiküszöbölésükre, különösen akkor, ha a referátum fordítása nyomtatásban is megjelenik.

Többnyelvű dokumentáció

Az információkban a következő elemek fordulnak elő:

- a) szerző
- b) cím eredeti nyelven
- c) cím lefordítva
- d) hivatkozások
- e) osztályozás
- f) kulcsszavak vagy deskriptorok
- g) referátum .

Az a), b), d) és e) elemek nemzetközies és függetlenek a nyelvtől. Ezek katalogizálási szabályoknak megfelelően kezelendők.

A kulcsszavak vagy deskriptorok megfelelő szótár segítségével géppel fordíthatók. Különböző nemzetközi rendszerek többnyelvű szótárai ez idő szerint már több szakterületen eredményesen használhatók és a deskriptorok szintjén lehetővé teszik a többnyelvű dokumentációs rendszerek hatékony alkalmazását mind az adattárakban, mind a referáló lapokban.

Az adattárak esetében a felhasználónak módjában áll, hogy az általa választott nyelven fogalmazza meg a kérdéseket, ami számos előnnyel jár mind a költségek, mind a pontosság szempontjából is. A kinyomtatott válaszban a cím és a referátum eredeti nyelven jelenik meg a felhasználó által választott nyelven megadott deskriptorokkal együtt.

A referáló lapok esetében többé-kevésbé elfogadható megoldások léteznek a rendelkezésre álló anyagi lehetőségeknek megfelelően.

Franciaországban jelenleg két megoldást tanulmányoznak. Az egyik a *kötött mondat szerkezetű referátumok* gépi fordítása. Előnye, hogy a gépi program egyszerűbb az általánosan használt gépi fordításénál, és csak néhány szerkezet használatára korlátozódik. A fordítás elvileg hibátlan, de csak a gyakorlat mutatja meg, hogy ez a kötettség megfelel-e az információ céljainak.

A másik megoldás a *kötetlen formájú referátumok* gépi fordítása. Előnye, hogy a legkülönbözőbb források-

1. táblázat

Egynyelvű, kétnyelvű és háromnyelvű információs rendszerek
költségszerkezetének összehasonlítása

Műveletek	Egynyelvű rendszer	Kétnyelvű rendszer		Háromnyelvű rendszer	
		deszkriptorok	referátumok	deszkriptorok	referátumok
		gépi fordítása			
Primer dokumentumok	9	18	18	27	27
Katalogizálás	4	6	5	7	5
Referálás	23	46	23	69	23
Indexelés	21	21	21	21	21
Gépi input	19	28	19	37	19
Gépi feldolgozás	3	10	30	14	42
Kiadás	21	42	42	63	63
Összesen	100	171	158	238	200
Költségszint kooperáció nélkül ¹⁾	100	200	200	300	300
Megtakarítás	0	14,5%	21%	21%	33%

¹⁾ Kétnyelvű rendszer helyett két egynyelvű, háromnyelvű rendszer helyett három egynyelvű rendszert tekintetbe véve.

ből származó referátumok minden előzetes átdolgozás nélkül felhasználhatók, ugyanakkor azonban hátránya az ismertett hibák felmerülésének lehetősége. A hibák előreláthatóan mintegy 2 év alatt kiküszöbölhetők, a stílus nehézsége azonban megmarad.

Költségek

A közös összefogásból eredő adattárak és a munkamegosztás feltétlenül költségmegtakarítással jár. Ennek elemzéséhez a többnyelvű információs rendszer két változata, nevezetesen a) a *deszkriptorok*, b) a *referátumok* gépi fordítását végző rendszer szolgál alapul.

A költségeket egyszerűbb százalékos megoszlásukban vizsgálni, mintsem a külső körülményekkel együtt változó tényleges összeget. Eszerint a referálólap kiadásának költsége a következőképpen oszlik meg:

primer dokumentumok beszerzése és kezelése	9%
katalogizálás	4%
referálás	23%
indexelés	21%
gépi adatbevitel és tárolás	19%
gépi feldolgozás	3%
kiadási költségek	21%
	100%

Aszerint, hogy 2 vagy 3 egynyelvű rendszert veszünk alapul, a költségek megduplázódnak vagy megháromszorozódnak. Kérdés, hogy egy kétnyelvű, illetve háromnyelvű rendszerben – mely 2, illetve 3 egynyelvű rendszert foglal magába – hogyan alakulnak ezek a költségek, számításba véve az említett a) és b) változatot.

A *primer dokumentumok beszerzése* tekintetében megtakarítás nem jelentkezik, az egyes költségek összegeződnak, feltételezve, hogy a többnyelvű rendszer beszerzi az eredeti rendszerek valamennyi dokumentumát.

Bár a *katalogizálás költségei* a nyelvek számától független, a 4% a többnyelvű rendszerekben azért emelkedik, mert a szabályokat szigorúbban kell alkalmazni és szükség van a munkamegosztás kérdéseinek rendszeres megvitatására. Az a) változat esetén az egyes rendszerek kiegészítő összefoglaló referátumokat is készítenek a központi adattár számára, a b) változatban azonban erre nincs szükség, ezért itt a költségek alacsonyabbak.

Ugyanezért megkétszereződik, illetve megháromszorozódik a *referálás költsége* az a) változatban, míg a b) változatban azonos marad.

A 21% *indexelési költség* mindegyik esetben azonos.

A *gépi input* az a) változat szerint 2 nyelvű rendszerben kétszer, 3 nyelvű rendszerben háromszor történik. Tekintve azonban, hogy a 19%-ból a referátumok inputja

csupán 9%, a maradék 10% a többi költség, tehát csupán a 9% duplázódik, illetve háromszorozódik meg. A költségek növekedése tehát

a kétnyelvű rendszerben $9 \cdot 2 + 10 = 28\%$,
a háromnyelvű rendszerben $9 \cdot 3 + 10 = 37\%$

A b) változatban a költség változatlan marad.

A gépi feldolgozás bonyolultabb lesz, ezért költsége is emelkedik.

A referálólapon előállítási (kiadási) költsége megkétszereződik, illetve megháromszorozódik, tekintve, hogy azt két, illetve három nyelven kell megjelentetni.

Következtetések

A táblázatból látható, hogy a költségmegtakarítás hosszú távon igen jelentős lehet. A megtakarításnak hozzávetőlegesen a fele további fejlesztésre fordítandó, így a tényleges megtakarítás mintegy felére csökken. Megjegyzendő, hogy a gépi feldolgozás költsége $3 \cdot 3 = 9\%$ helyett akár 42%-ra, azaz több mint a négyszeresére is nőhet. Ebből következik, hogy a tudományos és műszaki dokumentáció a közeljövőben hajtóereje lesz a gépi fordítás fejlesztésének.

/D'OLIER, J. H.: *Multilingualism in scientific and technical documentation = International Forum on Information and Documentation*, 2. köt. 4. sz. 1977. p. 20–24./

(Dezső Zsigmondné)



TÁJÉKOZTATÁS GÉPESÍTÉS

A CNRS reprográfiai szolgálata

Párizsban a CNRS (*Centre National de la Recherche Scientifique = Országos Tudományos Kutatási Központ*) reprográfiai üzemét első ízben 1967-ben láttam. Ebben az időben már korszerű gépi berendezésekkel, jól felszerelten működött és sikerrel alkalmazták az akkor még teljesen újnak számító mikrofilmlap technikát. Ezen a területen mint úttörők komoly eredményeket értek el; az általuk tervezett és kivitelezett felvevőgépekkel dokumentációs célra, nagyüzemi szinten, szabvány méretű 105 x 148 mm-es mikrofilmlapokat állítottak elő.

1971-ben az egész intézmény új, modern épületbe költözött, ahol a reprográfiai üzem is megfelelő elhelyezést kapott. A régebbi tagolt, elkülönített munkahelyek helyett áttekinthető, üvegfallal elválasztott helyiségeket rendeztek be. Időközben a laboratórium kitűnő vezetője J. J. Bastardie magasabb beosztásba került, az új vezető D. Goulard nagy szakértelemmel vezeti tovább az üzemet.

A CNRS Reprográfiai Szolgáltatának munkája jelenleg 4 fő részre tagozódik:

- mikrofilmtekercs készítés,
- felvételezés mikrofilmlapra,
- fotomásolatok készítése,
- nyomdaüzem.

A fentiekén kívül természetesen számos más fotomunkát is végeznek a lehetőségeiken belül. Az állandó létszámuk 50 fő, akik heti 40 órát, azaz 5 x 8 órát dolgoznak.

Az intézményhez 13 ezer időszakos kiadvány érkezik, ezekből 10–11 ezret dolgoznak fel. Ezekből a megrendelők részére foto- vagy mikrofilmmásolatokat tudnak ké-

szíteni. Ezenkívül kongresszusi kiadványokból és könyvtárközi kölcsönzés útján (esetenként külföldről) beszerzett mikrofilmekről is másolnak megrendelés esetén. Könyvek feldolgozásával egyáltalán nem foglalkoznak.

A saját folyóiratállományukon kívül egyetemi műszaki könyvtárakból is szolgáltatnak mikrofilmmásolatokat. Ebből a célból hetenként kétszer felkeresnek 6 nagy egyetemi műszaki könyvtárt, és megrendelések alapján a folyóiratokból kifényképezik a szükséges anyagot.

Évente kb. 8 millió oldalt másolnak, ennek 25%-át mikrofilm és mikrofilmlap formájában szolgáltatják, 75%-át pedig papírmásolatként. A mikrofilmlap szolgáltatás az évek folyamán fokozatosan emelkedik, míg a tekerescsfilmből, azaz filmcsik megrendelés csökkenő tendenciát mutat.

A napi megrendelések száma eléggé ingadozó: általában 1000 és 2000 tétel között mozog. Ennek ellenére az átfutási időt maximum 4–5 nap között tartják.

A szerzői jogra való tekintettel a dokumentumokból mindig csak egy másolatot vagy mikrofilmet készítenek és 3 hónapnál frissebb anyagot egyáltalán nem felvételeznek. Egyes publikációk szerzői jogát megveszik (pl. doktori disszertációk) és szolgáltatják, elsősorban mikrofilmlap formájában. Ugyanebben a formában dolgozzák fel az Unesco jelentéseket és több más hasonló dokumentumot is.

A tekerescsfilmből készítő részleg 5 felvevőgéppel dolgozik, ezekből négy 35 mm, egy pedig 16 mm szélességű filmre készíti a felvételeket. A 35 mm-es felvevők közül az egyik Kodak MRD 2 típusú, melyből nálunk is több üzemel, a másik három francia Debie gyártmányú. A 16 mm-es felvevőgép Filmaster típusú.

A mikrofilmezendő dokumentumok (főleg folyóiratok) a raktárból futószalagon érkeznek – a müncheni