

Centro Internacional de Agricultura Tropical

SEMINARIO


CIAT
64773
 COLECCION HISTORICA

HUNGSPAN: El Diccionario como Base de
 Datos y de Terminologia

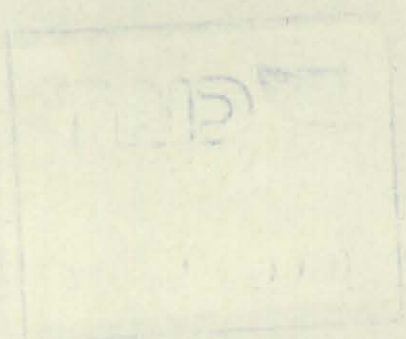

 BIBLIOTECA
 11 FEB. 1991
 4791


 COLECCION HISTORICA

11385

M. Isabel Bolton
 Encargada de Traducción
 por Computador
 Unidad de Publicación
 CIAT

52749



LIBRARY
ACETICA
1971

CONTENIDO

Resumen

Introducción

Breve descripción de ENGSPAN y sus programas hermanos

Uso de ENGSPAN en CIAT

Estructura de los diccionarios

Importancia de la actualización de los diccionarios

Operaciones de actualización

Unidades

El sistema como base de datos y de terminología

1) Vocabularios especializados

2) Microglosarios. Ejemplo de su uso

3) Vocabularios codificados según su fuente
terminológica

4) Vocabularios codificados según su confiabilidad

Perspectivas de ENGSPAN

Bibliografía

El material usado para la preparación de este borrador se tomó de:

1. Leon, Marjorie y Lee Ann Schwartz. "Integrated Development of English-Spanish Machine Translation: From Pilot to Full Operational Capability." Technical Report of Grant DPE-5543-G-86-3048-00 from the U.S. Agency for International Development. Washington, D.C.: Pan American Health Organization, Octubre 1986.
2. Muriel Vasconcellos. The Contribution of Machine Translation: Present Status and Future Perspectives." In: Microcomputer Applications in Education and Training for Developing Countries, proceedings of Second Symposium, Microcomputers in Education and Training (Cuernavaca, Mexico, 4-7 November 1985). Boulder and London: Westview Press. 1987. pp. 269-281.
3. Vasconcellos, Muriel y Marjorie Leon. SPANAM and ENGSPAN: Machine translation at the Pan American Health Organization. Computational Linguistics 11 (2/3):122, 136, 1985.

Resumen

Se describe brevemente el programa de traducción por computador ENGSPAN y los programas de traducción a otros idiomas derivados del mismo algoritmo. Se presenta una síntesis del trabajo de actualización del Diccionario realizado hasta el momento en CIAT, con ejemplos gráficos para explicar el proceso de actualización en sí: flexibilidad permitida, codificación, vocabularios específicos (Bases de Datos o Términos por instituciones, disciplinas, usuarios, etc.), operaciones de adición, cambio y eliminación de palabras individuales o de unidades en ambas partes del diccionario (inglés y español). Se explican los siguientes conceptos del trabajo con el diccionario: unidades de sustitución (simples, largas y demoradas), de análisis, de transferencia (para objeto, código, complemento, sujeto, modificativo, palabra principal, términos unidos por guiones, contexto hacia la derecha del término, contexto hacia la izquierda del término) y unidades de transferencia basadas en unidades de sustitución. Se tratan brevemente las perspectivas de este tipo de programa a la luz de los nuevos adelantos en la ciencia de sistematización.

Introducción

Con los vertiginosos avances de las ciencias de la informática y la sistematización y la creciente importancia de la traducción como herramienta de transferencia de conocimientos y tecnologías, los sistemas de traducción computarizada ganan importancia y se perfeccionan día a día. En este trabajo se explicarán algunos aspectos del sistema de traducción ENGSPAN, en forma necesariamente superficial e introductoria, dado lo escaso del tiempo y la complejidad del tema. Los aspectos tratados son: breve descripción del programa ENGSPAN y programas hermanos, incluyendo una explicación del proceso de traducción; uso de este programa en CIAT; estructura de los diccionarios e importancia de su actualización; operaciones de actualización; diversas bases de datos que contiene el sistema; y, por último, perspectivas de esta poderosa herramienta

Breve descripción de ENGSPAN y programas hermanos.

ENGSPAN (que significa inglés a español) es uno de los dos sistemas internos de traducción por computador (IBM mainframe 4341) desarrollados por la Organización Panamericana de la Salud (OPS), oficina regional para la Organización Mundial de la Salud (OMS), en Washington, D.C.

El otro es SPANAM (español a inglés americano). En la OPS, el trabajo en traducción automática se inició en 1976, usando el idioma español como idioma fuente. Entre 1976 y 1978 se montó la arquitectura del sistema, con un considerable apoyo de software, y se desarrolló el algoritmo básico para la traducción. La fase operacional se inició en 1980 cuando se logró implementar el análisis morfológico de los verbos y solucionar el problema presentado por el insumo de textos. Esto se logró instalando una interfase entre el mainframe IBM y las instalaciones Wang (Wang OIS/140) de procesamiento de la palabra. ENGSPAN se tornó plenamente operativo en agosto de 1985.

En ENGSPAN se ejecuta un algoritmo de traducción contra los dos voluminosos diccionarios del sistema. El diccionario inglés tiene aproximadamente 60.000 entradas, con sus correspondientes acepciones en el diccionario español. La velocidad de traducción es de aproximadamente 700 palabras por minuto, tiempo horario, y de 172.800 ppm, tiempo CPU. El sistema funciona con un parámetro de 215 K y sus dos diccionarios son archivos VSAM montados en discos permanentes, con un tamaño de 9 MB cada uno. El programa de traducción, escrito en PL/I, es muy complejo y comprende tres etapas principales: 1) análisis gramatical del texto en inglés, 2) transferencia estructural y lexicológica y 3) síntesis en español.

Para el análisis del texto en inglés se usa un formalismo llamado "red de transición aumentada" (Augmented Transition Network, o ATN)* mediante el cual se analiza la oración total con base en 175 tipos de información sintáctica y semántica, la cual se basa en los registros correspondientes a cada léxico. Esta misma información se usa para la transferencia estructural y lexicológica (ver ejemplo en la Ilustración 3). En la etapa de síntesis en español, se usan 101 criterios adicionales para generar estructuras que no existen en inglés, como por ejemplo, ciertos usos del subjuntivo.

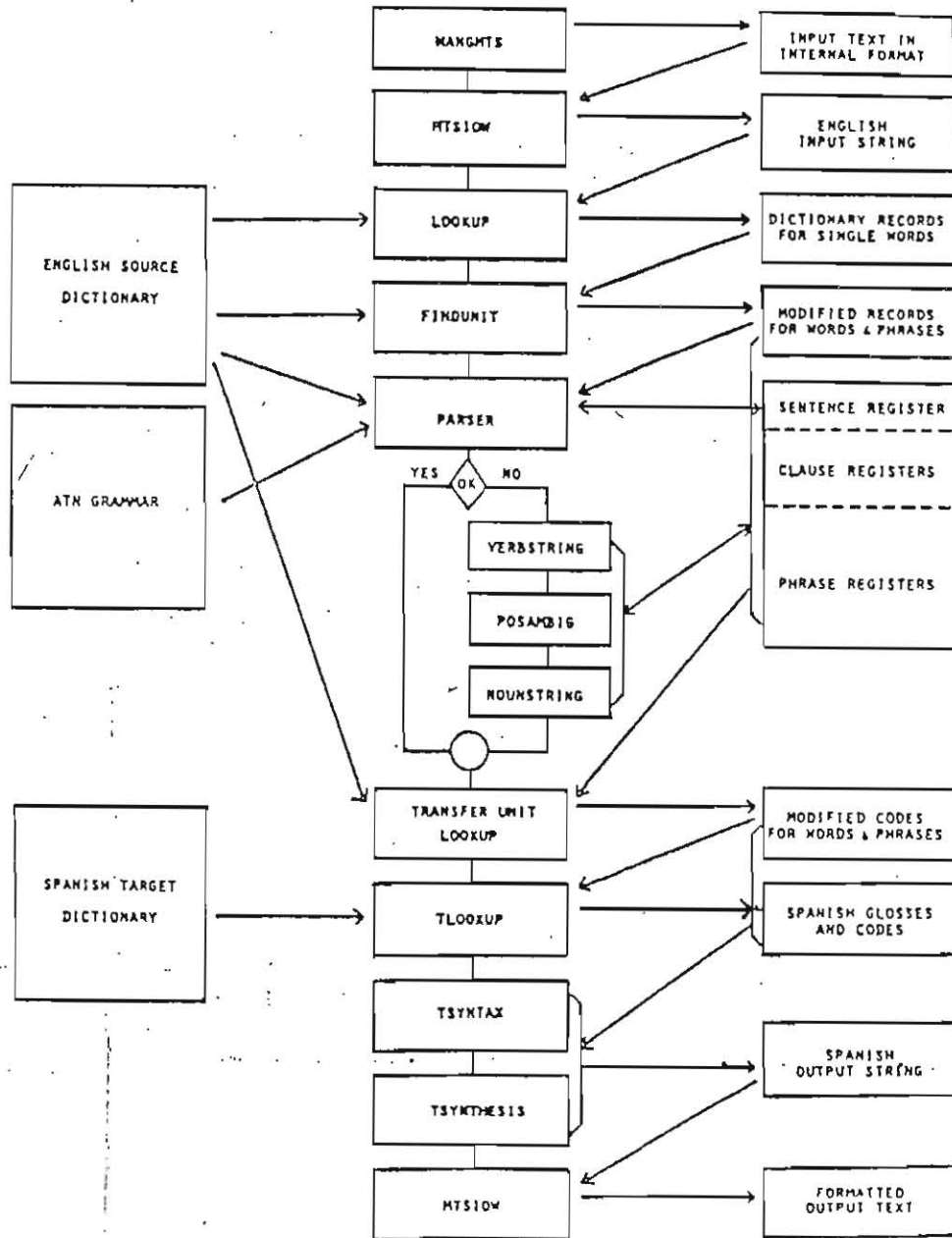
Con los sistemas de análisis del idioma inglés, de transferencia, y de síntesis del idioma español ya desarrollados para ENGSPAN, el trabajo está muy avanzado para la traducción automática del inglés al portugués, al francés y posiblemente a otros idiomas romance.

A continuación veremos una representación esquemática del proceso de traducción:

* La rutina ATN se deriva del trabajo de Winograd, Woods y Kaplan (ver Bibliografía)

Ilustración No. 1

Representación esquemática del proceso de traducción



Uso de ENGSPAN en CIAT

Mediante un compromiso con el GCIAI, el programa ENGSPAN se trajo a CIAT a mediados de 1987 y se hicieron las adaptaciones necesarias para que funcionara con el procesador de palabras Displaywrite en IBM (4361 LK5 6MB memoria real).

La utilización de ENGSPAN en DW4 implica varios pasos. Los documentos para traducción (escritos en DW4 o en cualquier otro sistema y luego convertidos en archivos ASCII) se transfieren con el software FTCMS (DCA file transfer) que opera con el Sistema IBM VM/CMS en el mainframe. Todo el proceso se hace en batch y el operario o usuario del programa solamente tiene que responder los menus que aparecen en la pantalla de su PC. El programa produce tres archivos: el SIDE-by-SIDE para uso exclusivo de la persona que actualiza los diccionarios, el SPANISH para revisión o post-edición, y el LISTING que contiene las estadísticas del trabajo realizado por el programa, para uso del personal de Sistemas.

En CIAT, el Servicio de Traducción por Computador se presta a los Programas y Unidades de CIAT. Desde que el programa llegó al Centro se han traducido aproximadamente 10.000 páginas de diversos tipos de documentos (informes anuales,

monografías, folletos, artículos de revistas, documentos de trabajo, etc.) sobre yuca, frijol, pastos y forrajes, arroz, estadística, economía, recursos genéticos, temas administrativos, actas de juntas, etc. También se han recibido trabajos de otras instituciones, como CIP, IRRI, ISNAR, y del CGIAR.

Los diccionarios

Estructura de los diccionarios. La estructura de los diccionarios en inglés y en español es la misma. Cada registro consiste de 161 bytes, agrupados en campos como se indica en la Ilustración No. 4. Nunca se llenan todos los campos del registro para ninguna palabra o unidad. La configuración de los campos depende de la parte (s) de la oración y de otros parámetros (por ejemplo, un verbo llenará diferentes campos que un sustantivo).

Importancia de la actualización de los diccionarios. Se han realizado aproximadamente 3500 operaciones de actualización en la copia de los diccionarios de ENGSPAN que posee CIA. Desgraciadamente, esta cifra es muy baja, considerando los extensos vocabularios de terminología agrícola que existen, los cuales crecen continuamente.

El perfeccionamiento de los diccionarios es de gran importancia para el mejoramiento de ENGSPAN como una herramienta efectiva de traducción y para satisfacer la necesidad de contar con vocabularios específicos (por ejemplo, para instituciones, programas, etc.). En este contexto, el sistema de traducción se podría considerar idealmente como un proceso interactivo permanente que involucra el ambiente institucional (las instituciones, hospedantes que suministran la terminología a incluir en los diccionarios), el ambiente de programación del sistema (las personas que pueden mejorar las capacidades del sistema), y el sistema mismo (los diccionarios como una base de datos en continua expansión y las capacidades de traducción del sistema).

Además, en la situación ideal arriba descrita, podemos afirmar que, en general, cualquier conclusión con respecto a la calidad del producto de ENGSPAN (i.e., traducciones producidas) y su edición tienen que ser necesariamente--e idealmente--de corta duración, ya que los diccionarios (y el sistema mismo) han sido diseñados para cambiar y mejorar continuamente.

Entre las características más apasionantes de la última versión del programa (en CIAT trabajamos con una versión anterior) se cuentan los nuevos tipos de unidades de

transferencia (ver Ilustración No. 6) que controlan el contexto anterior, palabras de prueba que se conectan con la palabra en cuestión, exploran la oración para determinar la presencia de una palabra específica o de una palabra con un conjunto específico de características semánticas, examinan el ambiente sintáctico para distinguir entre "count nouns" y "bulk nouns" (sustantivos que se pueden contar y sustantivos que se miden en bulto), etc. Otras características nuevas incluyen el mejor manejo del artículo definido, de las referencias bibliográficas (instrucciones para que no se traduzcan los nombres propios) y más opciones para el manejo de la voz pasiva (M. Leon, comunicación personal, 1990).

Operaciones de actualización. Para comprender mejor el alto grado de flexibilidad que ENGSPAN permite para el perfeccionamiento de sus dos diccionarios, revisemos brevemente los tres tipos básicos de operaciones de actualización (y su codificación) que se pueden realizar en los diccionarios: adición, cambio y eliminación:

Ilustración No. 2

Adición:

1. AES (Add English Source)
2. AST (Add Spanish Target)
3. ABD (Add Both Dictionaries)

Cambio:

1. CES (Change English Source)
2. CST (Change Spanish Target)
3. CBD (Change Both Dictionaries)

Eliminación:

1. DES (Delete English Source)
2. DST (Delete Spanish Target)
3. DBD (Delete Both Dictionaries)

Las siguientes instrucciones son para agregar SING al diccionario fuente, MUJER al diccionario meta, y MAN, con su correspondiente significado en español HOMBRE, a los dos diccionarios:

AES SING LEX=nnnnnn POS=V TNS=STEM CODA=1OBJ,2OBJ
AST MUJER LEX=nnnnnn POS=N GEN=F NUM=S
ABD MAN POS=N NUM=S FT=HUM,CONC,COUNT
HOMBRE POS=N GEN=M NUM=S

El diccionario puede contener muchos significados
alternativos para la misma palabra. Estos significados
corresponden a diferentes partes de la oración, áreas de
expresión, colocación de las palabras, ambiente sintáctico y
contexto semántico. Ejemplos:

Ilustración No. 3

Parts of speech

Plant (noun)	planta
	fábrica
Plant (verb)	plantar
	sembrar

Discourse type

nursery (hospital)	sala de recién nacidos
nursery (general)	habitación de los niños
nursery (agriculture)	vivero

Collocations

nursery school	escuela de párvulos
day nursery	guardería infantil
nursery rhyme	poema infantil

Syntactic environment

grow (no obj)	crecer
grow (direct obj)	cultivar
grow + adj	ponerse
grow + old	envejecerse
advise + inf	recomendar
advise + clause	informar
advise + on	asesorar

Semantic context

advise + hum obj	asesorar
advise + non-hum obj	recomendar

Ilustración No. 4

El Diccionario

ITEM	MEMBERS POS	GEN NUM TMS	LEXICAL HD	ITEM
__CODA=PREP				
018432001900 ADVICE ON	00		990009000000	
ADVISABILITY VOCAB=USA	N	S	077205000000	CONVENIENCIA VOCAB=USA
ADVISABLE CODA=TOINF, IMPCL FEATURES=DCL, MORE	A		070740000000	ACONSEJABLE
ADVISE CODA=10JJ, DCL, PREP, OBJTO, OBJPR, 10CL SURJ=+HUM 10BJ=+HUM OH=P IF=1	V	STEM	015127000000	ASESOR CLASS=AR TYPE=MIENTO CODA=IMPCL A=P SOBRE= AL=P
			015127000001	RECOMEND CLASS=AR
			015127000002	INFORM CLASS=AR TYPE=INF
015127 (01-CODA) ADVISE (1-CODA) CODA=TOINF	V	PRES	990100000000	
015127 (10-CODA) ADVISE (1-CODA) CODA=DCL	V	PRES	990200000000	
015127 (10-CODA) 001900 ADVISE (1-CODA) ON CODA=PREP	V	PRES	990009000000	
015127 (10-OJJ)--- ADVISE (1-OJJ) FEATURES=HUM, COLL	NUT V	PRES	990100000000	
ADVISER FEATURES=CONC, COUNT, ANIM, HUM	N	S	007003000000	ASESOR A=P SOBRE= AL=P
ADVISOR FEATURES=CONC, COUNT, HUM	N	S	074100000000	ASESOR
ADVISORY SOURCE=20 SU=2	A		056626000000	CONSULTIVO SOURCE=20
156606015712 ADVISORY SERVICE	V		200196000000	ASESORIA
ADVOCACY	N	S	074644000000	DEFENSA

Como ejemplo adicional de la flexibilidad del sistema, consideremos las posiciones 60-61 del registro del diccionario, las cuales corresponden a la parte principal de la oración, POS (ver ilustración anterior, primera línea):

Ilustración No. 5

Part of speech	CODE (mnemonic)	60-61 (numeric)
Preposition	PREP	01
Coordinate conjunction	CONJ	02
Subordinate conjunction	SUBC	03
Auxiliary	AUX	04
NCT		07
THAN and AS		09
Noun	N	11
Verb	V	12
Adjective	A	13
Adverb	ADV	14
Predeterminer		21
Determiner		22
Adjective premodifier		23
Ordinal number		32
Quantifier		33
Cardinal number		34
Proper name		38
Personal title		39
Subject pronoun		41
Subject and object pronoun		42
Object pronoun		43
Reflexive pronoun		45
Modifier of adjective and/or adverb		51
Modifier of numerative		52
MUCH		53
Prefix		75
Hyphen or slash		76
Punctuation		80-90

Durante el proceso de traducción, las palabras no halladas en el diccionario (o las palabras mal escritas en el documento original) reciben una asignación de POS que concuerda con su morfología. Algunas veces, las palabras no halladas en el diccionario hacen que el análisis gramatical se efectúe sólo parcialmente o no se efectúe. En este caso, hay de todas maneras una traducción, la cual se basa en un análisis de cada frase, más que en las acciones realizadas por el ATN.

Unidades. Cuando es necesario incluir un registro en el diccionario fuente (inglés) que contiene más de una palabra (p. ej., in spite of, web blight) el codificador del diccionario puede optar por utilizar un registro múltiple, o sea de varias palabras. Este tipo de registro se llama unidad. Existen los siguientes tipos de unidades:

Ilustración No. 6

Unidades

1. Substitution units

simple substitution units (SSU), contain 1-5 words

long substitution units, contain 1 or more nested SSU

delayed substitution units, to add verb-adjunct pairs

2. Analysis units

3. Transfer units

To test an object (nn-OBJ)

To test coda (nn-CODA)

To test a complement (nn-COMP)

To test a subject (nn-SUBJ)

Triggered by a modifier (nn-MOD)

Triggered by a Head (nn-HEAD)

To test a hyphenate compound (nn-HYP1)

To test the right-hand context (nn-CONR)

To test the left-hand context (nn-CONL)

4. Transfer units based on substitution units

Ilustración No. 6 (Ejemplos)

Simple SU:	IN SPITE OF	A PESAR DE
Long Su:	(800123) FUND FOR POPULATION ACTIVITY FONDO DE LAS NACIONES UNIDAS PARA ACTIVIDADES EN MATERIA DE POBLACION	
Delayed SU:	TAKE IN	INGERIR
Analysis unit:	OVERSEAS POST	
Transfer unit:	CALL (nn-CODA) FOR	

El sistema como base de datos y de terminología

Entre las bases de datos que contiene el sistema se cuentan

1) los vocabularios de un determinado usuario, o vocabularios especializados; 2) los microglosarios; 3) los vocabularios codificados según su fuente terminológica; 4) los vocabularios codificados según su confiabilidad.

1) Vocabularios especializados. La posición cero del registro del diccionario se divide en ocho campos de bit, cada uno de los cuales contiene una señal que identifica el vocabulario especial de un usuario específico. Los bits se "prenden" al especificar el código VOC (ver la Ilustración No. 3, arriba). Los vocabularios especializados, que se pueden recuperar en pantalla o en papel usando la identificación VOC correspondiente, son:

Ilustración No. 7

VOC=WHO	World Health Organization
VOC=AGRIS	AGROVOC Thesaurus
VOC=CIAT	Centro Internacional de Agricultura Tropical
VOC=USA	U.S. Agency for International Development
VOC=UN	United Nations

2) Microglosarios. Además, hay varios microglosarios que proporcionan una traducción alternativa, específica al área, para un término que ya posee una traducción más general en el diccionario, como por ejemplo leaflet, que significa hojita en botánica y folleto en impresión. Los microglosarios son los siguientes:

Ilustración No. 8

Microglosarios	Código
Initial caps	MICRO=1
All caps	MICRO=2
Super Medical (Research)	MICRO=3
Legal and Financial	MICRO=4
Sanitary Engineering and Environmental Health	MICRO=5
CGIAR Agriculture	MICRO=6
General Agriculture	MICRO=7
Primary Health Care	MICRO=8
Correspondence	MICRO=9
Computer Science (AID)	MICRO=10
United Nations	MICRO=11
Reserved for Atomic Energy	MICRO=12

Ejemplo del uso de los microglosarios. Las siguientes operaciones de actualización tienen el efecto de agregar un nuevo significado (en el área de la agricultura) a la palabra NURSERY, que ya se encuentra en el diccionario:

Ilustración No. 9

CST SALA DE RECIE/N NACIDDS LEX=073615 MICRO=7

AST VIVERO LEX=073165 45-46=07

3) Vocabularios codificados según su fuente terminológica.

También es posible marcar la fuente terminológica (y lograr su posterior recuperación en listados o en pantalla) de una palabra usando los siguientes códigos:

Ilustración No. 10

PAHD	00
Medical	10
General	20
AGROVOC Thesaurus	70
AID	80

4) Vocabularios codificados según su confiabilidad.

Además, el producto del sistema (o sea, la traducción) contiene indicaciones de los significados equivalentes que se pueden hallar en los microglosarios especializados y de los términos que tienen un alto puntaje de "confiabilidad".

Con esto, queremos decir que una palabra en cuya aceptabilidad podemos confiar recibe el código RCODE. El mayor puntaje, o sea RCODE=5 significa que se trata de un término estandarizado a nivel intergubernamental; el menor puntaje, o sea RCODE=0 significa que no se ha tomado ninguna decisión con respecto al estado terminológico del término.

Recapitulemos ahora los diferentes conceptos expresados hasta ahora sobre el uso del diccionario como base de datos y de terminología:

1. Los mismos diccionarios son una base de datos (bilingüe) en continua expansión.
2. El sistema contiene vocabularios especializados (VOC=WHO, AGIS, CIAT, USA, ONU).
3. Además, el sistema proporciona microglosarios especializados de términos alternativos para áreas específicas (Investigación Médica; Materias Legales y Financieras; Ingeniería Sanitaria y Salud Ambiental; Agricultura de GCIAI; Agricultura General; Atención Primaria de Salud; Correspondencia; Ciencia de Computación (AID); Naciones Unidas; y Energía Atómica).
4. El sistema proporciona vocabularios codificados según su fuente terminológica.
5. El sistema posee también una base de términos calificados según su "confiabilidad".

6. El uso del sistema en microcomputador también proporciona acceso a bases de datos lexicológicas, en ambos extremos del sistema, el de entrada o insumo y el de salida o producto.

Por otra parte, quisiera llamar la atención del lector al hecho de que las bases de datos corrientes son un fait accompli y, en consecuencia, algo que debemos aceptar y acerca de lo cual no podemos hacer nada (Dewey, 1982). Entre una base de datos corriente y el documento que suscita la búsqueda hay una persona, un operario. Los diccionarios de ENGSPAN, por el contrario, interactúan con cada palabra de cada documento que entra (es digerido) por el sistema y proporcionan retroalimentación constante, y cada vez más efectiva, si se atienden adecuadamente. Al procesar (digerir) ENGSPAN todo el documento (y con los nuevos avances logrados en el reconocimiento de caracteres ópticos, prácticamente cualquier libro, documento, carta, manuscrito, nuevo o viejo, puede ser así procesado) produce algo a cambio de cada palabra, término, unidad, código, signo, etc. que el documento contiene. Qué produce exactamente depende de cómo se han cargado sus bases de datos. Consideremos la diferente cantidad de información que ofrecen estos dos productos:

Ilustración No. 11

monogragraph

monografía

bonavist bean

arbejón Lablab purpureus (L.)
Sweet (Dolichos lablab L., D.
purpureus L., (L. niger Med.,
L. vulgaris Savi)

Las implicaciones de todo esto se pueden dejar a la imaginación y al vertiginoso desarrollo de las ciencias de sistematización e informática, especialmente ahora con las perspectivas apasionantes de las búsquedas de textos completos, un avance que permite poner en línea libros completos, y donde cada palabra puede ser sujeto de una búsqueda (Dewey, 1982).

Perspectivas de ENGSPAN

El futuro de la traducción por computador (ENGSPAN y SPANAM) en la OPS está tomando tres direcciones principales: 1) la implementación de estos dos sistemas en otras instituciones poseedoras de mainframes, especialmente en aquellas desde donde se pueda servir a la comunidad de naciones del Tercer

Mundo; 2) el desarrollo de pares de idiomas adicionales; y 3) la adaptación de los diferentes sistemas al ambiente de microcomputador. En este documento trataremos con cierto detalle sólo este último punto.

En el sistema ENGSPAN el microcomputador se usa (como procesador de palabras) tanto en el extremo de entrada o insumo como en el extremo de salida o de producto. Hasta cierto punto, el microcomputador también proporciona acceso a bases de datos lexicográficas, la base propia del traductor y las grandes bases de datos para las cuales el microcomputador puede servir como terminal en-línea. Esta última tecnología, la cual se explota actualmente con SUSANNAH, un ambiente de microcomputador desarrollado específicamente para traductores por la Universidad de Saarbrücken, tiende a maximizarse a medida que se amplían las redes de telecomunicaciones y se hace más disponible la tecnología de acceso en pantalla.

Actualmente, con las mayores capacidades de memoria desarrolladas (ahora la memoria se puede medir en GIGAs, o sea miles de millones de bytes) es técnicamente factible situar un sistema grande de traducción automática en un microcomputador, aunque es posible que tengan que sacrificarse ciertas operaciones complejas. Además, el análisis gramatical del idioma fuente, dependiendo de la

tecnología, puede ocupar gran parte de la CPU y desacelerar considerablemente un sistema. Por otra parte, los idiomas de programación que ofrecen la mejor utilización de las capacidades del microcomputador, por ejemplo "C", no son aquellos usados originalmente para escribir las versiones de mainframe. En el caso de SPANAM y ENGSPAN, las versiones originales se escribieron en PL/I. La OPS planea escribir nuevamente todo el sistema en "C", con todas las ventajas que esto implica (M. Leon, comunicación personal).

Bibliografía

1. Agricultural Terms, as Used in the Bibliography of Agriculture from Data Provided by the National Agricultural Library. 2d. Ed. (Phoenix) Dryx Press, 1978. Revision de 1976 Thesaurus of Agricultural Terms. 122 p.
2. AGROVOC Thesaurus. Edición en inglés, Ene. 1981. [Roma, AGRIS], 1981. 507 p.
3. Dewey, Harry. The AGRICOLA database> a textbook. Beltsville, Md. National Agricultural Library. II.A.21, 1982.
4. Kaplan, Ronald M. A General Syntactic Parser. In: Rustin, Randall, 1973: 193-241.
5. Leon, Marjorie y Lee Ann Schwartz. "Integrated Development of English-Spanish Machine Translation: From Pilot to Full Operational Capability." Technical Report of Grant DPE-5543-G-86-3048-00 from the U.S. Agency for International Development. Washington, D.C.; Pan American Health Organization, Octubre 1986.
6. Vasconcellos, Muriel. The Contribution of Machine Translation: Present Status and Future Perspectives." In: Microcomputer Applications in Education and Training (Cuernavaca, Mexico, 4-7 November 1985). Boulder and London: Westview Press. 1987: 269-281.
7. Vasconcellos, Muriel y Marjorie Leon. SPANAM and ENGSPAN: Machine translation at the PanAmerican Health Organization. Computational Linguistics 11 (2/3):122, 136 1985.
8. Winograd, Terry. Language as a Cognitive Process. Vol. 1 Syntax. Addison-Wesley. Reading, Massachusetts. 1983.
9. Woods, W.A. Augmented Transition Network for Natural Language Analysis. Report No. CS-1 to the National Science Foundation, Cambridge, Massachusetts, 1969.