

El Lenguaje y la Informática. Dependencias mutuas

Antonio Vaquero

**Escuela Superior de Informatica
Universidad Complutense de Madrid
E-mail: avaquero@eucmax.sim.ucm.es**

1. Introducción

El Lenguaje y la Informática están influidos mutuamente de muchas formas.

La Informática penetra en nuestra sociedad cada vez mas intensamente. Los hábitos y el lenguaje se modifican continuamente por este proceso.

El lenguaje es el legado cultural mas importante de una comunidad. En su conformación influyen muchos factores, pero actualmente hay dos muy importantes: la jerga científica y técnica (particularmente la informática) y el uso de las computadoras. Los efectos de esta influencia son mucho mas intensos en el mundo educativo que en cualquier otro sector de la sociedad. Hay que prestar atención no Solo al lenguaje de comunicación entre profesor y alumno, sino también al "software" educativo.

En primer lugar es conveniente analizar como se incrementa el lenguaje desde la Ciencia y la Tecnología, particularmente en el caso de la lengua española, para intentar racionalizar ese proceso, al que se le debe exigir la característica de homogeneidad o generalidad dentro de una misma comunidad lingüística.

Para particularizar al campo de la Informática es preciso estudiar antes las relaciones entre ésta, la Educación y el Lenguaje.

Este análisis se efectúa tanto en la comunicación entre profesor y alumno como entre éste y las computadoras. Este estudio proporciona una perspectiva desde la que se clarifican las características exigibles a la Informática de consumo.

Naturalmente se particularizan toda esas consideraciones al caso del español.

2. Relaciones entre el lenguaje y la informática

El caso de la Informática es particular dentro del proceso de la influencia de la Ciencia en el Lenguaje. Pero es un caso singular por dos motivos. El primero es el propio dinamismo de la Informática, cuyo "corpus" de conocimientos se agiganta cada día. El segundo es la repercusión mutua entre Lenguaje e Informática. Si la Informática expande el lenguaje, el lenguaje influye poderosamente en los sistemas informáticos. Estos efectos han de ser estudiados cuidadosamente para comprender la importancia del problema que estamos tratando. Vamos a considerarlos a la luz de la influencia del sistema educativo.

3. La enseñanza y el lenguaje

En términos educativos el usuario es el alumno. El mundo educativo ha de ser muy sensible al correcto uso de la lengua por su influencia en la formación y su efecto multiplicativo. El uso de términos incorrectos, así como el incorrecto uso de los términos, tiene virtudes deformadoras, mas que formadoras, del buen criterio que se pretende ir creando en los hábitos intelectivos de la persona que se esta formando.

Particularmente la Enseñanza de la Informática y de las materias relacionadas con la misma ha de ser extraordinariamente cuidadosa en este sentido. Hay que tener en cuenta además que la Enseñanza de la Informática debe ser escrupulosa en el respeto a la lengua

materna de la población infantil.

Instituciones de ámbito universal, como UNESCO e IFIP prestan atención a este asunto. Así se ha elaborado el documento "Informatics for Secondary Education" (van Weert, 1994). La traducción al castellano de ese importante documento ha de hacerse con sumo cuidado.

El efecto del profesorado y de las interfaces persona sistema informático en el buen uso del lenguaje va a ser determinante, no solo en la cultura informática y la normalización de los términos técnicos, sino también en el dominio de la propia lengua. La Informática Educativa debe ser portaestandarte de esta preocupación.

Un nuevo problema que tiene que afrontar nuestra sociedad en relación con la Informática y el Lenguaje es que los usuarios están siendo fuertemente influidos por las interfaces. Analicemos este problema, que tiene dos caras: una relacionada con la entrada al sistema, o sea, la forma en que se responde, se pregunta o se requiere algo por parte del usuario y otra relacionada con la salida, o sea la forma en que el sistema comunica algo. Vamos a comenzar por la primera.

4. Influencia de la informática en el lenguaje del usuario

La ubicuidad de la Informática se manifiesta en los hábitos de la sociedad y muy particularmente en el lenguaje. No me refiero aquí al lenguaje técnico, que debe enriquecer nuestro diccionario, aunque falta mucho por hacer en este sentido, sino al lenguaje de la calle, que se corrompe con giros y palabras, incorporando, la mayor parte de

las veces inadecuadamente, la metáfora de la computadora. En los EE.UU. de Norteamérica se llama "Technobable" a esa forma de hablar, que constituye toda una lengua en los ambientes de los "hackers", pero que se ha introducido también en el lenguaje de la calle. (Barry, 1991).

A pesar de la sofisticación creciente de las interfaces para hacer cada vez mas natural el dialogo persona maquina, éstas imponen de hecho una restricción en el uso de las posibilidades de expresión del lenguaje. Con el uso intensivo de las computadoras existe un riesgo de empobrecimiento del lenguaje que, de hecho, esta siendo observado como un fenómeno que ocurre en sociedades mas avanzadas. De esta manera aparece un problema educativo consistente en como ha de ser el sistema integral de enseñanza que asegure, contemplando un uso racional de los sistemas informáticos, el desarrollo de habilidades lingüísticas que permitan explotar la potencia expresiva del idioma enteramente.

5. Una informática para cada tipo de destinatario

Por lo que respecta a la segunda cara del dialogo persona sistema, el problema global consiste en que el "software" incorpora ineludiblemente características que son particulares de una cultura y un lenguaje dados. El equipo productor de un paquete de "software" para un conjunto de aplicaciones esta formado por un grupo de personas con una cultura y lengua comunes y el producto que desarrollan esta pensado

para ser utilizado por unos usuarios con unas características culturales y lingüísticas determinadas, concretas y especificadas. Si el usuario tiene esas características, entre todas la más importante la lengua, los mensajes del sistema serán muy naturalmente comprensibles si el "software" tiene alta calidad.

Así pues se podría hablar de una "Informática original", Informática hecha para determinados grupos de usuarios de una determinada comunidad x parlante. Es obvio que esa condición de "originalidad" debería ser exigible a todos los productos informáticos. Y no solo al "software", sino también a todos los elementos visibles de la interfaz. Piénsese, refiriéndonos a nuestro idioma concretamente, en el tan traído y llevado problema de la eñe, por ejemplo.

El efecto de destinar un producto informático "original" a grupos de otra comunidad lingüística puede ser muy dañino. Para los usuarios pertenecientes a otras culturas y con una lengua materna distinta de la original aparecen problemas de comprensión de los mensajes que no desaparecen simplemente con la mera traducción de los mismos. Y ello aunque la traducción se buena, caso que no es en absoluto frecuente.

Si limitamos el fenómeno en cuestión a la españolización de los paquetes "originales" en inglés, se pueden observar un sinnúmero de tremendas incorrecciones lingüísticas.

Como ya anticipamos, no se trata solo de que los mensajes estén bien traducidos, sino sobre todo de que las especificaciones de los requisitos del sistema, desde el comienzo del

ciclo de desarrollo del "software", tengan en cuenta las características de cada grupo diferente de usuarios, entre las que hay que destacar la lengua materna de los mismos. ¿Por qué comprar y usar lo que se ha hecho para otros?

Así pues, propugnamos no una Informática "original" en inglés, ni una Informática españolizada, aunque esté españolizada por técnicos hispanoparlantes, caso insólito hasta hoy, sino una Informática "original" en español.

6. El caso de la informática educativa

En el caso del "software" educativo, el fenómeno tiene, además otras connotaciones que pueden resultar peligrosas para la formación de los estudiantes, e incluso contrarias a los objetivos de las políticas educativas. Analicemos estas peculiaridades.

La calidad del "courseware", como "software" que es, es susceptible de ser medida como éste. Pero tiene características específicas. Para aplicarle los criterios de cumplimiento de objetivos, hay que tener en cuenta el entorno, o sea, ¿es el destinatario el mismo previsto por los autores? (Forse, Th., 1989). En el caso del mundo educativo este desenfoque puede ser nefasto. Cuando el "courseware" es foráneo, no solo puede haber problemas de lenguaje ("software" españolizado), con toda su importancia, sino problemas pedagógicos por las diferentes mentalidades y formación de los destinatarios. Esto sucede en muchas ocasiones y no se puede impedir más que produciendo nuestro propio "software".

7. Aplicaciones muy sensibles a la lengua

Inciendo en la influencia del lenguaje sobre el "software" educativo, es conveniente observar que el grado de influencia puede depender de las aplicaciones particulares. Vamos a exponer a continuación un conjunto de ellas en que ese grado es muy alto.

Hay aplicaciones educativas en las que se ve muy claramente la necesidad de producir "courseware" y herramientas "originales". son aquellos sectores del "software" educativo en los que el problema del idioma es esencial, como el caso de la enseñanza de idiomas asistida por computadora, campo que esta dando muy buenos frutos. Dentro de cada comunidad lingüística puede haber dos clases de "courseware" de enseñanza de su lengua:

a) Para los miembros de la comunidad (aprendizaje de la ortografía, lectura comprensiva, etc.)

b) Para enseñanza del idioma a extranjeros.

Proyectos del tipo b) para enseñanza del inglés se analizan en Johnson (1986).

En lengua francesa esta el proyecto EMILE (Depover, 1991), que tiene ya desarrollada la enseñanza del francés para lusoparlantes con técnicas multimedia interactivas.

Pensando en nuestra comunidad y teniendo en cuenta la expansión que actualmente tiene el idioma español, nos aflora con fuerza una demanda: ¿Para cuando un proyecto de "courseware original" de enseñanza del español para extranjeros? ¿Esperaremos a que los

angloparlantes produzcan los cursos de español para ellos como destinatarios? ¿Esperaremos que los franceses hagan lo mismo?

¿Tendremos los iberoamericanos que pasar por el sonrojo de comprar los cursos de español para lusoparlantes o los de portugués para hispanohablantes producidos fuera de la comunidad iberoamericana? De esta parte de la tarta de lo que se ha dado en llamar "industrias de la lengua", ¿tampoco vamos a probar bocado?

8. Presencia de la lengua en el mundo

Pero, aun con ser muy importante lo que podemos perder como industria informática o, mucho más, como cultura y potencialidad informática, aún es más importante lo que podemos perder como lenguaje y como cultura lingüística en confrontación, queramos o no, con otras culturas en el contexto universal. Tenemos, pues, la obligación de atender con cuidado a la contribución que la Informática debe hacer al español "general".

Otras comunidades lingüísticas prestan gran atención al desarrollo de la Informática y su enseñanza en su lengua. P. e. en abril de 1996 se celebrará el "5: Encuentro Francófono sobre la Didáctica de la Informática".

Por otra parte, la atención que al español se le presta desde países no hispanoparlantes es una muestra de la importancia de nuestra lengua. Mencionaremos un par de ejemplos europeos.

Existe un proyecto alemán de televisión digital interactiva, que

desarrolla la Deutsch Telekom. Dentro de ese proyecto la emisora DW TV emite un programa internacional en inglés y en español. En español se titula "La voz de Alemania".

La cadena británica BBC emite un curso de castellano con el título "suecos" con el que pretende enseñar no solo la lengua sino también la cultura española, en un sentido amplio, tanto en su ámbito histórico como geográfico.

9. ¿Qué podemos hacer los informáticos?

En primer lugar, conocer lo que ya se ha hecho, aunque sea insuficiente. En Vaquero, A. (1992) se hace una breve glosa histórica de los trabajos previos a 1992.

En el mismo trabajo se analiza la terminología informática en español y se propugnan unas pautas para usar el lenguaje correctamente. También se justifica la necesidad de crear un Diccionario multilingüe y un vocabulario de términos técnicos de Informática en español. Mejor que Vocabulario se debería decir Glosario, pues es preciso definir el concepto y no solo traducir el término inglés (Recomendación del panel "Terminología and Concepts involved in Computer Communications and Teleteaching" Teleteaching'93 IFIP. Trondheim, 25 agosto 1993).

En el mismo sentido se exponen las bases para un proyecto hispanoamericano para cubrir esas lagunas.

El tiempo transcurrido desde entonces hace sentir aún más ese tipo de carencias.

Quizá bastaría echar una ojeada a lo que se hace en comunidades más aventajadas e intentar hacer lo propio en la nuestra. Vamos a tomar un caso significativo en esta línea de actividad: las Bases de Datos Léxicas.

10. Bases de datos léxicas

Las bases de datos léxicas constituyen una extensión del concepto de diccionario y resultan muy adecuadas para ser utilizadas en muchas aplicaciones que involucran la comunicación en lenguaje natural. En inglés existe la base de datos léxica Wordnet. Wordnet (Miller, 1993) es una base de datos on line de libre disposición que contiene información fundamentalmente léxica, extremada de forma semiautomática a partir de diccionarios. Wordnet es un proyecto desarrollado en el Cognitive Science Laboratory de la Universidad de Princeton y se basa en ideas tomadas del campo de la psicolingüística relativas a la organización de la memoria léxica humana.

El sistema Wordnet se compone de archivos lexicográficos, código para convertirlos en una base de datos, rutinas de búsqueda e interfaces para visualizar la información de la base de datos.

En Wordnet se ha intentado construir un nuevo tipo de "diccionario": un diccionario en el que se puedan realizar búsquedas conceptuales, no solo alfabéticas.

De ahí que Wordnet se parezca más a un thesaurus que a un diccionario.

La gran diferencia entre un

diccionario estándar y Wordnet reside en la organización que Wordnet hace del léxico. El contenido de Wordnet está clasificado en cuatro categorías correspondientes a categorías sintácticas diferentes (nombres, verbos, adjetivos y adverbios). La ventaja de esta categorización sintáctica es que las diferencias fundamentales en la organización semántica de estas categorías sintácticas se ven claramente y se pueden explotar de modo sistemático. A cambio hay que pagar un precio: la existencia de información redundante, cosa que no ocurre en los diccionarios convencionales (por ejemplo, palabras como "set" y "back" pueden pertenecer a varias categorías diferentes).

Dentro de cada una de las divisiones del léxico, la información se ha estructurado mediante significado y, para cada término, es posible obtener otros tipos de información que solo se puede conseguir en diccionarios especializados, como los enciclopédicos o los de sinónimos y antónimos.

Una aplicación de Wordnet a la construcción de sistemas de "software" se ha realizado (Fernández Chamizo, 1995) en el Departamento de Informática y Automática de la Universidad Complutense de Madrid. La aplicación consiste en la recupe-

ración de componentes de "software" en Programación Orientada a Objetos.

En general Wordnet tiene una virtualidad de aplicación muy extensa a la recuperación de información de cualquier tipo mediante consultas en lenguaje natural.

Hay aplicaciones en las que los usuarios formulan su consulta en inglés de forma natural, bien por ser su lengua habitual de comunicación o bien porque es de uso profesional común, como es el caso en equipos de programadores profesionales.

Pero hay muchas aplicaciones con usuarios que necesitan hacer las consultas en su propia lengua, como son prácticamente todas las aplicaciones educativas, aunque no solamente éstas. Por ello es necesario emprender cuanto antes trabajos encaminados a producir una base de datos léxica del español, por lo que concierne a la comunidad informática hispanohablante.

11. Recomendación general

Para aumentar el nivel, tanto cualitativa como cuantitativamente, de la producción informática en nuestra comunidad lingüística es necesario tener en cuenta al usuario hispanoparlante. En general en

cada país hay que tener en cuenta las particularidades del usuario y muy especialmente la lengua. Los países que han tenido en cuenta lo expresado en esta recomendación han tenido mayor desarrollo informático, como lo expresan Baeza Yates et al. (1995). ■

REFERENCIAS

- Baeza Yates, R.A., Fuller, D.A., Pino, J.A. y Goodman, S.E. (1995) "Computing in Chile: The jaguar of the Pacific Rim?" *Communications of the ACM*. Septiembre 1995, pp. 23-28.
- Barry, John A. (1991) "Technobabble". The MIT Press.
- Fernández Chamizo, C. et al. (1995) "Promoting Software Reuse Through Explicit Knowledge Representation". *EPIA'95, Madeira*, octubre.
- Johnson, M. & Osguthorpe, R.T. (1986) "Computer Based Foreign Language Instruction: Improving Student Attitudes". *Educational Technology*, Spther. pp. 24-28.
- Miller, G.A. et al. (1993) "Five papers on Wordnet", *Cognitive Sc. Lab. Princeton Un.*, CSL Report 43.
- Vaquero, A. (1992) "Informática, Educación y Lenguaje" *Revista Enseñanza y Tecnología (ADIE)* n: 7, pp. 18 a 31.
- van Weert, t. (editor) (1994) "Informatics for Secondary Education" UNESCO