

## DEL OBJETO DE LA LINGÜÍSTICA Y DE SU OBSERVACION \*

### I

Al ser la lengua un instrumento de comunicación se acostumbra estudiarla con un doble enfoque: como instrumento y según la información transmitida. Pocas veces se distingue claramente el contenido del código; éste suele estudiarse a través del prisma del contenido y eso cuando el estudio no se reduce tan sólo al contenido; los trabajos sobre el código en sí son escasísimos.

Diversas razones, a menudo accesorias, explican la falta de interés por el código o la predilección de que goza el contenido. La aparente facilidad que tiene el niño para aprender una o varias lenguas y la total inconsciencia del adulto en la práctica cotidiana podrían no ser ajenas a esta orientación. La satisfacción de la meta alcanzada en el acto de comunicación o el empeño por llegar a ella hacen comprensiblemente que pase a segundo plano la reflexión sobre los medios empleados. Tampoco es excepcional que el prestigio —semántico, metafísico, estético u otro— del producto dé lugar a cierto desdén por el instrumento utilizado en su elaboración.

La confusión del código y del contenido —del significante y del significado, del *designandum* y del designado, del referente y del referido, de la gramática y de la semántica, etc.— puede tener también causas más profundas. En la comunicación, tanto al producir como al percibir, ambos componentes se interfieren a menudo pro-

---

\* Ponencia leída en el XV Simposio de la Sociedad.

fundamente, restringiéndose y respaldándose a la vez. Sin el uno, el otro no existe: el mensaje sólo se transmite a través del canal del código preconcebido, y éste puede ser indescifrable si no le preceden conocimientos previos, ajenos al mismo código. El pensamiento, individual o colectivo, somete el código a sus necesidades, y éste le imprime sus huellas. Cuando uno de ellos viene a faltar o falla, voluntaria o involuntariamente, el otro queda sin control o se disgrega. En tales condiciones no debe sorprender que se explique uno con arreglo a otro, o mediante el otro, y que de este modo se tome a veces uno por otro.

Las interferencias entre formas y contenidos se han explotado a menudo, más o menos a sabiendas y con diversos fines. En los refranes, dichos y otros giros, la sabiduría popular identifica y fija el mensaje mediante la consonancia de las formas. La didáctica intenta autenticar los conceptos basándolos en el análisis —etimología exacta o falsa— de las palabras. Lo mismo ocurre en poesía cuando se da por hecho que el sonido sugiere el contenido o preside al génesis o la estructura del poema.

Desacoplar las asociaciones convencionales es también frecuente y puede surtir efectos variables. Éstos pueden provenir de intervenciones sobre el referente o sobre el referido. Ciertas modificaciones pueden ser introducidas voluntariamente con fines precisos —como en literatura, por ejemplo—, o resultar de cambios incontrolados en la utilización del código —en caso de afasia, por ejemplo—, o de un nuevo reconocimiento de la realidad referida, en el espacio o en el tiempo. En cada caso hay límites que no se pueden rebasar so pena de incomunicación. Los lenguajes técnicos, las jergas, la poesía, el lenguaje infantil, las producciones patológicas proporcionan con frecuencia ejemplos de desacoplamiento de naturaleza y grado variados.

\* \* \*

Saber si el referente es anterior al referido o se desprende de él, cuál de ellos es primario, esencial o causal, es una cuestión que no se plantea al tratarse de un adulto que habla normalmente su primera lengua. Cronológicamente no es seguro que tanto en lo que respecta a la humanidad como al individuo, el problema esté bien planteado.

Que la expresión sea la expresión de un contenido no lleva a que el contenido determine la organización de la expresión. Y lo contrario tampoco es cierto. Una lesión cerebral puede privar a un individuo de la palabra durante algún tiempo, sin que cesen forzosamente su comprensión o su concepción; el esquizofrénico puede hacer un discurso demente sin errores de gramática. La lengua poética más enrevesada puede expresar sentimientos o sensaciones de lo más corriente. Las asociaciones lingüísticas incongruentes de los niños tienen sólo la originalidad semántica que los adultos están dispuestos a atribuirles. Cada cual ha podido experimentar que reconoce lo que no puede enunciar y más de uno habrá asombrado nombrando lo que es incapaz de concebir.

Las restricciones del código no determinan el orden semántico, pero contribuyen a limitarlo, cercarlo u orientarlo en cuanto a las relaciones lógicas o aspectos expresables, por ejemplo, tanto en la mente del escritor (locutor) como en la del lector (oyente). Las proscripciones semánticas tampoco determinan el orden gramatical, pero imponen una selección entre las posibilidades que el código ofrece, en lo que a las combinaciones léxicas se refiere, por ejemplo. La selección depende de la cultura de la época, la sociedad y el individuo; es el criterio tan relativo de la verdad. Importa no confundir los condicionamientos semánticos —locales, temporales y subjetivos— con las reglas gramaticales; también importa no explicar el mundo según el modo de designarlo en uno u otro lugar.

En esta confrontación la lengua se encuentra sin duda en posición débil. Los vocablos, clases morfológicas, relatores sintácticos que el código tiene a su disposición no abrazan, ni de lejos, la infinita variedad de objetos o nociones, aspectos o relaciones existentes o potenciales. En la generación como en el análisis se confirman cotidianamente la dificultad de encontrar la palabra exacta y la insuficiencia, vanidad y vaciedad de tantas fórmulas. La desagregación sintáctica y la neología en literatura o la sustitución de la gramática y el vocabulario naturales por lenguajes artificiales o lógicos, en el proceso de datos por ordenador, por ejemplo, son otras tantas tentativas de remediar la inadecuación del referente lingüístico para expresar lo referido o de paliar su distancia.

Es cierto que la lengua puede revelarse también como un instrumento dúctil y moldeable a las necesidades de la comunicación en

la conversación como en literatura. No es menos cierto que se trata de una armonía excepcional, parcial e ilusoria, cuya evidencia no se impone a todos de igual modo ni en cualquier momento. La movilidad del lenguaje resulta irrisoria si se la compara con la que se percibe en la realidad sensible y la que anima el pensamiento. Para la mayoría el aprendizaje precoz y ciego y las convenciones condenan al código a una paralización relativa, mientras que el mundo sensible y pensando —en tanto no se exprese—, se diversifica sin traba alguna.

Es cierto también que la relación arbitraria y convencional entre significante y significado puede mostrar estabilidad en el espacio y en el tiempo, y que la relación más estable es la que afecta al nivel sintáctico: los cambios son lentos en el tiempo y las desviaciones apenas toleradas en un estadio de lengua determinado. Esta relación parece tanto más estable cuanto que las relaciones entre vocablos y nociones o cosas se revelan estar sometidas a modificaciones frecuentes y rápidas en el transcurso del tiempo y experimentar notables variaciones sociolectales. La diferencia de estabilidad puede hacer creer en una mayor cohesión o incluso identidad entre significante y significado por lo que a las relaciones atañe, encubriendo así el conflicto latente entre la expresión y lo expresado, un escollo inevitable y casi insalvable, y la piedra angular de todo intento de expresión estética o didáctica. Quien estudia la comunicación lingüística halla a cada paso indicios del conflicto: en la morfosintaxis se juegan el porvenir de una lengua, la perennidad de un texto literario o la pertinencia de un tratado filosófico; y en ese mismo plano ha tenido que inventarse el hombre nuevos lenguajes, artificiales, para poder abordar mejor el mundo físico, incluido a veces el estudio de las lenguas naturales.

Tampoco hay que olvidar que toda organización gramatical o semántica es relativa. Una y otra se manifiestan como realizaciones individuales que no siempre son coherentes en sí ni coinciden forzosamente con las organizaciones ajenas. El carácter fragmentario, en muchos casos desequilibrado, de estas realizaciones retira al conjunto de la praxis lingüística y al universo semántico de cada individuo, sociedad o cultura lo que de sistemático pudieran tener. El sistema se relaja o se desquicia en las realizaciones dispares de cada individuo, en la colectividad y en el tiempo, y tal disparidad contri-

buye a debilitar las ya azarosas relaciones entre la expresión y lo expresado.

A esto se añade, para terminar, que la lengua es el producto y el instrumento de un órgano humano, el cerebro, dotado de inmensas cualidades al tiempo que, como cualquier órgano, es naturalmente imperfecto en comparación con un sistema teóricamente ideal. La lengua no puede evitar compartir las características que son propias del órgano que la crea y no puede sino adaptarse a sus imperfecciones naturales para ser eficiente. Cargada de redundancias y de ambigüedades, con un desarrollo irregular, sometida unas veces a un sinnúmero de restricciones caprichosas y otras entregada a todas las licencias, una lengua natural dista mucho de alcanzar el grado óptimo en cuanto a la cantidad de información teóricamente transmisible. En las condiciones y para los objetivos acostumbrados de la comunicación humana los códigos inequívocos, perfectos en sí, como los que exige el cerebro electrónico, por ejemplo, no pueden convenir al cerebro humano que sin preparación o ayuda especial es incapaz de abarcar la totalidad de la información encerrada en cada signo o combinación de signos y organizarlos con total eficiencia. Si la comunicación lingüística, la transmisión de información, la articulación del contenido y de la expresión se considerasen como manifestaciones «orgánicas» —en el sentido literal de la palabra—, quizá los planteamientos al respecto no se formularían de la misma manera.

\* \* \*

Para estudiar la lengua, sus elementos constituyentes y la organización de estos mismos, importa aislar cada uno de ellos. Sólo se puede reconocer o simular, enseñar o transformar —explicar—, un fenómeno complejo si se le puede descomponer, o sea distinguir el papel de cada parte en el conjunto. Es la mejor garantía para no atribuir a las formas y a sus combinaciones lo que pertenece a la interpretación semántica, o, dicho de otro modo, es la condición para determinar y medir lo que sólo pertenece a las formas. Posteriormente esta distinción inicial respalda de manera óptima el indispensable estudio de las interferencias entre las formas y los efectos de sentido posibles.

Disecar, aislar, separar, sondear a fondo lo que la mayoría de las veces se hace inconscientemente es como un trabajo de laboratorio en cuyo beneficio redundan la utilización de un instrumental moderno. Los medios técnicos autorizan, a la vez que obligan, a eludir el sentido y a analizar o comprobar hasta los límites extremos los recursos del código. Argüir que un análisis con objetivos diferentes de los que se han previsto comúnmente —la comunicación— o con medios distintos de los habituales —los órganos humanos— es injustificado porque es artificial es tan absurdo como pretender que el análisis literario es incompatible con la lectura. Así vista la lingüística aplicada es a la lingüística lo que la estilística al análisis literario.

En la actualidad el instrumento de laboratorio por excelencia, a menudo imprescindible para los otros instrumentos también, es la computadora, sea cual sea su potencia. Posee todas las cualidades requeridas para proceder al trabajo de aislamiento progresivo de los elementos buscados en la complejidad de los datos y hasta sus carencias pueden resultar virtudes. Memoria virgen en un principio sólo puede proceder según instrucciones puntuales y ordenadas. En los datos suministrados sólo puede reconocer aquello para lo que previamente se la ha programado. En todos los aspectos la computadora es actualmente el instrumento más potente, preciso y objetivo para observar, medir, transformar y generar datos lingüísticos.

Dentro de esta línea no basta al lingüista que se hayan producido automáticamente resultados análogos a los de la mente humana, con un algoritmo de análisis sintáctico arborescente, por ejemplo, o la traducción por ordenador de una lengua a otra. El resultado no garantiza por fuerza que se haya simulado efectivamente el proceso lingüístico humano y menos aún el funcionamiento del código propiamente dicho. Las posibilidades, en más y en menos, de los sistemas lógicos o del cerebro electrónico son hasta tal punto diferentes de las del cerebro humano, que no se pueden conjeturar sino unidades, asociaciones y decisiones distintas en un orden diferente.

## II

No hay empresa intelectual que no pueda preciarse de algún resultado feliz por un motivo u otro. La trayectoria de una investigación

puede ser variada: rectilínea o sinuosa, prevista o inesperada, meritoria por la dirección emprendida o por el objetivo alcanzado, a veces simplemente perseguido o incluso a pesar de tal objetivo.

Todo objeto o noción, lo expresado como la expresión puede ser ideado de múltiples maneras. Es probable que todas y cada una de ellas no convengan en igual medida a cada objeto ni tampoco a cada uno de nosotros. En la manera de abordar y abarcar cada objeto u objetivo planeado se reflejan asimismo el momento, el lugar y la intención, con arreglo a los conocimientos ya adquiridos y a los ambicionados. La historia de la lingüística no es sino una larga serie de modos de aproximarse a la lengua, en la que el último es sólo el inicio de otros nuevos modos de proceder.

En ciencia importa que la marcha seguida sea conocida: que sea consciente y comunicada. Y sin embargo, no siempre sucede así por muy diversos motivos, voluntarios e imperdonables una veces, involuntarios y excusables otras. Los resultados de una investigación no reflejan forzosamente en su exposición el proceso y el método seguidos. Con respecto a ellos o a la teoría establecida éste puede parecer anecdótico. Las perspectivas generales en las que encaja un hallazgo pueden surgir claramente tan sólo más tarde. Etc.

Abunda lo escrito a propósito de la relación entre la teoría, la concepción general o la hipótesis de trabajo y la observación de los hechos. La historia nos enseña que más de una teoría no ha desembocado en resultados concretos, y que más de un hallazgo se ha abierto paso a pesar de las teorías vigentes; también nos muestra que hay hipótesis que han hallado su confirmación mucho tiempo después de haber sido formuladas. La experiencia demuestra, por fin, que una combinación acertada de reflexión teórica y observación pragmática principia una feliz conclusión.

En las Facultades de Letras, en las que se incluye tradicionalmente la Lingüística, se ha implantado una enseñanza en la que se concede el primer puesto a la teoría, cuando no a la historia de las teorías, a expensas de la observación, en la que se antepone el modelo al análisis, en la que el estudiante sólo aprende a reconocer lo que se conforma o no a la definición y a pasar de la regla a la aplicación. En los casos extremos, y no sin arrogancia a veces, los datos están tan seleccionados y escardados, tan manipulados o incluso silenciados que ya ni se intuye de qué la teoría podría ser la abstracción o la síntesis.

El razonamiento avanza según una lógica interna. El esquema concebido engendra clases y reglas, órdenes y niveles según evidencias que le son propias. De aquí se deriva una coherencia que otorga a la teoría visos de generalidad, universalidad, profundidad o elegancia que satisfacen a la mente, pero que no abarcan necesariamente la diversidad o la complejidad de la realidad. Algo así como artefactos que funcionaran perfectamente bien, pero en balde.

A este tipo de enseñanza se le puede hacer el reproche de ser cuando menos incompleta, en sentido único y, a nuestro modo de ver, a contrasentido de lo que sería preciso hacer en primer lugar o por lo menos paralelamente en la universidad: aprender a explorar de modo sistemático y crítico los hechos lingüísticos y a deducir la definición, la regla, la ley o la teoría de la observación activa de los datos y su organización. Tal enseñanza contribuye a impulsar el estudio autónomo a lo largo de toda la vida después de los años de formación, y es una condición previa a la experimentación, clave de la renovación y de la indagación científica.

\* \* \*

Los hechos lingüísticos pueden captarse de diversas maneras. Para unos será el conocimiento intuitivo o ajeno, solicitado conforme a las necesidades; otros prefieren consultar conjuntos de datos previamente establecidos, realizados independientemente de la investigación y reunidos según criterios variables. La elección operada al respecto no es ajena a las preferencias anteriormente comentadas: habitualmente la intuición goza del favor de los teóricos.

Las ventajas y los inconvenientes de ambas actitudes han sido evocados reiteradamente. Con frecuencia se arguye que sólo la intuición puede apurar todas las virtualidades de la lengua. Es olvidar que las posibilidades combinatorias de los lexemas son en todo caso subjetivas y, por lo tanto, sometidas a controversia; es también infravalorar la riqueza de los corpus, incluso cortos, en el campo morfosintáctico y desconocer los medios de que se dispone actualmente para la exploración exhaustiva de corpus extensos y variados. No es obvio que los ejemplos que pueden cosecharse, interrogándose a sí mismo o a otros, deben o pueden insertarse en un mismo sistema, efectivamente realizado o practicable. Se puede poner en duda si los

ejemplos producidos intencional y posteriormente para dar respuesta a una pregunta precisa tienen un valor probatorio tan grande como los que surgen independientemente de la cuestión. Las condiciones de empleo de tales ejemplos no pueden ser sino meramente imaginarias y su probabilidad de aparición desconocida. La misma argumentación es válida para todo lo producido por cualquier intervención (transformación, supresión, ampliación, sustitución, mutación, etcétera) operada o rechazada.

Recurrir a un corpus sólo se justifica plenamente si se trata de un corpus cerrado y examinado exhaustivamente. Ir a buscar donde sea aquello que la demostración reclama, abre paso a cualquier tipo de contraejemplo. El examen exhaustivo permite graduar la representatividad de cada ejemplo. Para ello es preciso que el corpus sea homogéneo, seguido —o que pueda dividirse en fragmentos igualmente homogéneos y seguidos—, y lo bastante extenso. Los resultados de la investigación son válidos únicamente con respecto al corpus examinado, pero reales, seguros y objetivos. Cualquier generalización, tanto si parte de un corpus como si resulta de la introspección, debe hacerse con las mayores reservas. Para fundamentarla basta multiplicar los corpus, variar sus orígenes en el tiempo y el espacio, en lo que atañe al registro sociolingüístico, género literario, etcétera. La observación de las variaciones que intervienen, incontrolables intuitivamente, favorece una mejor comprensión de la relatividad de los hechos.

Muchos de los corpus utilizados hoy en día se componen de fragmentos delimitados estadísticamente, sin sujeción a la organización interna (principio, fin, división) de los textos. De este modo queda garantizada la objetividad de la selección en los conjuntos elegidos. El inconveniente es que estas muestras truncadas no permiten estudiar los fenómenos lingüísticos en su contexto. Éstos no pueden ser aprehendidos sino dentro de su marco natural, que, en muchos casos, es el texto completo.

Sólo los corpus constituidos de textos seguidos y recogidos íntegramente permiten percibir las relaciones paradigmáticas que, al tiempo que la red de restricciones sintácticas, completan, precisan, fijan o explican la expresión. A la inversa de las relaciones sintácticas que no traspasan determinados límites espaciales o temporales, las relaciones paradigmáticas abarcan espacios o intervalos extensos

aunque variables, más o menos definidos y cuya evidencia y regularidad varía según el autor, y muy probablemente según el género o el registro lingüístico (lengua hablada o escrita, poesía o prosa, etc.). Difieren según el tema, las circunstancias, el punto de vista del autor a la hora de redactar, y las convenciones del género.

\* \* \*

La explotación de corpus como los descritos más arriba queda facilitada mediante el recurso a la cuantificación, ya que ésta proporciona bajo diversas formas, informaciones objetivas y comparables a otras obtenidas del mismo modo, permite representar sucintamente y dar a conocer rápidamente datos más numerosos que antes, más complejos, a veces completa o parcialmente insospechados por ser inconscientes, cualesquiera sean el espacio y el tiempo en que ocurren. Puesto que los datos cuantitativos expresados en cuadros o gráficos, pueden transformarse a discreción, a tientas, conforme a criterios alternados o modificados sistemática y gradualmente, según preguntas puntuales o a todo evento, pueden orientar y diversificar la investigación, sugiriendo hipótesis, un orden en la investigación o una distribución de los datos que emanan de los mismos datos. En el análisis constituyen un punto de partida y un hilo conductor imparciales y reveladores para aquel que ignore realmente, finja ignorar —en virtud de la duda sistemática, por ejemplo—, o crea saber de manera intuitiva o razonada y que aspire a proceder a una comprobación.

Una operación de cuantificación no podría ser llevada a cabo sin los medios adecuados: un aparato estadístico adaptado a los datos o las necesidades lingüísticas y un juego de programas que abra paso a la utilización de calculadoras electrónicas. Sólo la computadorización puede garantizar óptimamente la recolección de los datos y su elaboración estadística. De hecho el ordenador es la clave de la lingüística que preconizamos aquí.

El recorrido no carece de obstáculos para el lingüista. Aún hoy en día la estadística y la informática se incluyen raras veces en su formación. No siempre puede acceder al ordenador en las condiciones requeridas y los prejuicios que todavía reinan en muchas facultades de letras son numerosos.

Tanto por eso como de por sí conviene que la lengua no sea un mero pretexto de cálculos sin ton ni son que hacen poco caso de las entidades o funciones reales y naturales, o las tratan desconectadas de la realidad compleja de la que proceden. La cuantificación no exime al lingüista del análisis complementario cualitativo. La primera es tan sólo una introducción al segundo como también es cierto que estamos convencidos que éste sale siempre ganando al apoyarse en aquella.

### III

Hemos insistido más de una vez sobre la oportunidad, la necesidad incluso, de introducir la computadora en el estudio de la lengua tal como lo entendemos. Sabida es la importancia de este instrumento relativamente reciente en las ciencias exactas, aplicadas y en un gran número de disciplinas de las llamadas ciencias humanas. Esto no es menos cierto en Lingüística, o por lo menos no debería serlo ya. También es conocido el exceso con que unos adoran y otros vituperan. No está de más intentar de vez en cuando situar la computadora en su justo lugar, acabando con el mito de unos y el anatema de otros.

Hablar de la computadorización supone al tiempo abrir el expediente de la lingüística aplicada. La noción o término «lingüística aplicada» entraña dos actitudes u orientaciones muy diferentes entre sí. Se puede tratar de la aplicación de conocimientos lingüísticos en campos distintos como la enseñanza de lenguas extranjeras, la logopedia o la traducción automática, por ejemplo. Es cierto que al estar en contacto con la práctica, el lingüista puede aprender mucho, pero se tratará siempre de adquisiciones residuales. Por lo general, esto es lo que hoy se entiende por lingüística aplicada.

Otra orientación, mucho menos adoptada, menos conocida y admitida, y menos espectacular, es la que consiste en recurrir a medios técnicos, métodos y herramientas diversos para indagar cómo funciona la lengua y cómo se produce y se asimila. El laboratorio de fonética acústica era, hasta fechas recientes, el mejor ejemplo. Utilizar el ordenador como instrumento de observación del objeto lingüístico, tal como lo hemos definido anteriormente, es hacer lingüís-

tica aplicada en el segundo sentido. Y es una fuente de renovación fundamental.

Sin embargo ésta no ha sido la utilización que del ordenador se ha hecho en lingüística en los treinta últimos años, desde que al principio de los años cincuenta se intentó por primera vez una traducción automática. O cuando alguna vez ha sido así, se ha hecho de manera desperdigada y las aplicaciones han pasado inadvertidas tanto para el lingüista como para el público. En general son los grandes proyectos espectaculares, como la traducción automática, los que han monopolizado la atención bien que, en términos lingüísticos, son auténticos monstruos. En tales condiciones el ordenador es un instrumento de fabricación o simulación que debe producir a toda costa en lugar de ser un instrumento de investigación y de observación no siempre rentable a corto plazo: el lingüista debe alimentarlo, servirlo y no servirse de él, y a la lengua no le queda más remedio que plegarse a la máquina en vez de ser analizada con su ayuda. En la mayoría de los casos esto es lo que se entiende hoy por lingüística computacional en la que demasiadas intervenciones pragmáticas anulan o falsean las perspectivas lingüísticas.

Así pues, cada vez que se trate de computadorización, convendrá subrayar esta divergencia de intereses. Y hay que hacerlo tanto más cuanto la distinción no siempre queda claramente establecida o concedida, sobre todo si la trama se esconde, a veces solapadamente, detrás de una cortina de alta tecnicidad. Citaremos dos ejemplos extremos y admitidos, y por lo mismo transparentes al respecto.

Existen en el mundo, incluido para el español, un cierto número de programas de lematización o segmentación morfológicas, más o menos elaborados, con fines diversos: la traducción automática, por ejemplo, o más sencillamente con vistas a un estudio de vocabulario (literario u otro). Consisten, en líneas generales, en reconocer en las formas ocurrentes —las que se encuentran entre dos blancos— una base y afijos comparando los segmentos sistemática y sucesivamente separados con repertorios de bases y afijos, y en verificar las combinaciones autorizadas. En los programas más afinados el análisis de las formas que siguen inmediatamente a la forma examinada y a veces de las que preceden, permite resolver los casos en los que varias segmentaciones (lematizaciones) son posibles fuera de contexto. Ni que decir tiene que la máquina no reconoce los afijos, las bases y

sus combinaciones más que en la medida en que éstas se hallan previstas en los repertorios previamente establecidos. Este tipo de programa utiliza los conocimientos morfológicos tradicionales, siempre y cuando éstos convengan. Porque si no es así, se adapta sin más la morfología mediante la invención de morfemas ficticios o la supresión de otros que no son rentables en el proceso automático. El interés propiamente lingüístico de tales procedimientos, por muy apreciables o indispensables que puedan resultar desde otros puntos de vista, puede ponerse en duda.

De otra índole es el programa que no dispone ni de catálogo de morfemas ni de reglas de asociación específicas, y se desarrolla al margen de toda interpretación semántica o etimológica, suponiendo únicamente que las lenguas naturales funcionan de forma más o menos económica, es decir, contable. Si tal programa logra producir, como se ha podido comprobar, segmentaciones aceptables para el usuario común, todas las segmentaciones posibles o probables y no otras, si se comprueba, como se ha hecho, que el proceso automático puede repetirse para otros inventarios de la misma lengua y de otras lenguas emparentadas, sin modificación alguna, se habrá demostrado que más allá, o más acá, de la razón semántica o etimológica funciona un sistema de segmentación que pertenece en propio al código y justifica o facilita su explotación semántica en sincronía y su evolución diacrónica.

En la documentación automática el problema estriba en reducir automáticamente textos escritos en una lengua natural a una lengua documental semi-lógica o semi-natural, comprensible para los especialistas de la cuestión tratada. La sustitución de las relaciones sintácticas entre las palabras por relaciones lógicas es una de las etapas importantes. Suele basarse parcialmente en la explotación de las informaciones morfosintácticas presentes en el texto mediante un diccionario, como el descrito más arriba, pero se basa principalmente en informaciones extratextuales, preexistentes y resumidas en una «red de nociones» específica del tema. Es un sistema de relaciones paradigmáticas, establecidas de antemano entre los vocablos y jerarquizadas empíricamente. Dentro de una disciplina determinada, éstas se reputan «estables y universalmente aceptadas» y tienen precedencia sobre las relaciones gramaticales, de las que se dice que son «contingentes, incluso contestables, y que no son generalmente verdad

más que en el contexto particular». El análisis se ayuda con los morfemas instrumentales que están en el documento, pero los que determinan en definitiva e incluso corrigen la relación entre éstos son los índices de relación nocional que acompañan en el diccionario a los lexemas cualquiera que sea la expresión sintáctica en la lengua natural.

Con todo ello el ordenador no deja de ser un magnífico instrumento y la lingüística tiene el mayor interés en hacerse con él. Permite explorar sistemática y exhaustivamente, de modo reiterado y fragmentado bancos de datos inmensos, incluso infinitos a propósito de unidades infinitamente pequeñas de manera rápida, es decir, económica, y contar, clasificar y estructurar los resultados según un número igualmente infinito de criterios variados. Hace posible un análisis cuantitativo y a través de él representaciones gráficas que hacen aparecer los resultados de manera objetiva, precisa y diversificada. Al tiempo se vuelven a plantear nociones tales como documentación, contexto, representatividad, gramaticalidad, norma e intuición.

El ordenador obliga al lingüista a descomponer los procesos de análisis o de síntesis lingüísticos, a definir de manera unívoca los criterios y las unidades utilizados y a establecer un orden preciso de las decisiones sucesivas que deben ser tomadas. Revela por lo absurdo los fallos en definiciones, razonamientos y jerarquías, puesto que sólo puede ejecutar lo que ha sido formulado explícitamente y sólo puede clasificar, analizar, generar o transformar según criterios previamente suministrados y que pueden ser reconocidos en los datos tratados.

La computadora permite efectuar en masa transformaciones de orden lógico o matemático, reiterables y susceptibles de ser modificadas gradualmente, fragmentadas e interrumpidas, y representarlas de manera simbólica, numérica o gráfica, es decir, objetiva, precisa y variada. Contribuye a acusar la distancia entre la lingüística y las ciencias especulativas, con las que está aún emparentada únicamente por tradición, acercándola a las llamadas ciencias exactas y aplicadas. Incluso el contacto en el ámbito del laboratorio de cálculo universitario con la máquina y sus usuarios, programadores e investigadores de todas las disciplinas que recurren al mismo instrumento,

puede ser para el lingüista una experiencia enriquecedora desde muy diversos puntos de vista.

La computadora evidencia, por fin, la relatividad de las cosas. A causa de sus límites y exigencias, a través de sus resultados y fracasos, invita al lingüista a poner constantemente en duda las ideas recibidas, formular diferentemente las reglas convencionales, adoptar otros puntos de vista y multiplicar los enfoques de la materia lingüística. En un mundo maniatado por las capillas, los dogmas y las presuposiciones teóricas, reintroduce, y no es paradoja, la imaginación, esa cualidad tan esencial a la investigación creadora, siempre que se acompañe del más exigente rigor científico como es ineludiblemente el caso con el ordenador.

JOSSE DE KOCK