

El problema lógico de la adquisición lingüística desde la perspectiva de la Teoría de la Optimalidad: comparación con el modelo de Principios y Parámetros y algunas consideraciones críticas*

VÍCTOR M. LONGA
Universidade de Santiago de Compostela

I. INTRODUCCIÓN

La adquisición del lenguaje puede ser contemplada como una especie de *puzzle* donde no todas las piezas son diáfanos; de hecho, como es bien sabido, diferentes corrientes han postulado explicaciones muy divergentes, e incluso contrapuestas, para dar cuenta del proceso de adquisición. La visión de la Gramática Generativa (desde ahora, GG), sostiene que algunas piezas importantes de ese *puzzle* no necesitan ser aprendidas, al derivar directamente de la contribución de un estado inicial común a la especie.

El objetivo de este trabajo consiste en exponer la visión de la adquisición sostenida por la Teoría de la Optimalidad (en adelante, TO), reciente teoría generativa y, por tanto, postura que defiende igualmente que piezas importantes de ese *puzzle* son independientes de la experiencia. Se mostrará el planteamiento y la resolución del problema lógico de la adquisición lingüística según esta teoría. Al tiempo, compararé las premisas usadas por ella con las de su oponente, Principios y Parámetros (desde ahora, P&P), y efectuaré algunos comentarios críticos sobre la perspectiva de Optimalidad, afirmando que la resolución del problema lógico propugnada por ella quizás no sea tan óptima como sus defensores pretenden y que el cuadro trazado por P&P es más sencillo y más adecuado para las premisas de factibilidad para el aprendizaje (en el sentido de Wexler & Culicover 1980).

El trabajo se estructura de la siguiente manera: en el apartado 2 se enuncian, a grandes rasgos, las bases del por qué de la defensa de la existencia de un problema lógico y el modo en que se ha encarado en GG. El apartado 3 presenta de manera esquemática los principales rasgos de TO, mientras que el 4 expone la concepción del problema lógico en el seno de esta teoría. El apartado 5 efectúa algunos comentarios

* Una versión previa de este trabajo fue redactada durante una estancia como *visiting scholar* en el *Romance Linguistics and Literature Program* de UCLA, realizada en enero-febrero de 1999. La mencionada estancia ha sido posible gracias al apoyo económico de la Xunta de Galicia. Deseo agradecer la acogida del Programa, por medio de su *Chair*, Dominique Sportiche, y especialmente la de Carlos Otero. Algunas de las ideas se han beneficiado de discusiones con Nina Hyams.

críticos sobre diferentes aspectos centrales asumidos en ella. Finalmente, se sintetizan las principales conclusiones del presente trabajo.

2. EL PROBLEMA LÓGICO: SUS BASES Y SU RESOLUCIÓN EN EL MODELO P&P

La GG defiende una concepción biológica del lenguaje y de su adquisición. La naturaleza de esta habilidad se atribuye a una capacidad innata, parte de la dotación biológica de la especie. Entre las variadas razones para adoptar este enfoque innato, destaca el argumento de la pobreza de los datos o 'problema de Platón' en Chomsky 1986 (cfr. Wexler 1991), que apunta precisamente a que el aprendiz no dispone en algunas áreas de datos diáfanos (a veces, de ningún tipo de datos) al alcance como para llegar al estado estable, y sin embargo, lo consigue (cfr. Crain 1991, Crain & Nakayama 1987, Crain & Thornton 1998, Hornstein & Lightfoot 1981, Lightfoot 1982, 1991, entre otros). Esto es, existe un vacío entre los datos disponibles y la gramática del niño como entidad mental al final del proceso. Un aspecto importante en relación con lo señalado consiste en el tipo de evidencia que el aprendiz maneja: no se le informa de la gramaticalidad de todas y cada una de las secuencias de una lengua, ya que ésta constituye un conjunto infinito, lo que significa que sólo está a su alcance la evidencia positiva, no la negativa (cfr. Gold 1967, Wexler & Hamburger 1973, Baker 1979, Marcus 1993, entre otros). Esto tiene consecuencias importantes: las secuencias que recibe como datos primarios constituyen sólo un subconjunto del total de secuencias gramaticales. Pero a pesar de que muchas secuencias gramaticales no son encontradas por el niño, éste converge de manera rápida hacia la gramática estable. No sirve decir a este respecto que el niño no enuncia secuencias agramaticales simplemente porque no las oye, ya que hay muchas buenas que tampoco oye pero que dice. En la misma manera, tampoco sirve la generalización como recurso único para salvar la ausencia de evidencia negativa, ya que ésta no trabaja en bastantes casos, produciendo hipotéticos resultados incorrectos que, sin embargo, no son enunciados por el niño (cfr. ejemplos en Piatelli-Palmarini 1989, Uriagereka 1998: 3-4, y especialmente Crain & Thornton 1998 entre otros).

Puesto que la información de los datos se queda bastante más corta que la gramática alcanzada, parte de ese conocimiento lingüístico se defiende como innato, no aprendido. Así, se postula un mecanismo de adquisición que permita dar cuenta de la factibilidad de aprendizaje¹, *Language Acquisition Device* (LAD), o simplemente, Gramática Universal (GU), que es el que permite dar cuenta de la unificación del proceso, ofreciendo una hipótesis sobre la rapidez y uniformidad de adquisición independientemente de la lengua, así como aunando el resto de las razones que sustentan

¹ Esto es importante, puesto que existen dos problemas en esta corriente: mientras en la perspectiva enunciada, correspondiente al problema lógico, se busca factibilidad para el aprendizaje, la óptica del aprendizaje en sí (desarrollo y conformación de las gramáticas) alude al llamado problema psicológico, o físico desde Chomsky (1980).

la visión innata (período crítico, 'paradoja del aprendizaje', propiedad específica a la especie, etc.)².

El problema lógico, o problema de la proyección (Peters 1972, Baker 1979), formaliza precisamente que a partir de una experiencia limitada, accesible al niño de manera asistemática, no ordenada, etc., puede aprender un objeto tan complejo como una lengua con gran rapidez y de manera inconsciente. Ese problema se ha abordado mediante la especificación de las gramáticas posibles dada una teoría lingüística que proyecta esas gramáticas. Sin embargo, la tarea debe ser acotada: ya que el niño aprende con rapidez y facilidad (independientemente de la lengua), esto significa que muchas de las posibilidades en número y clase desde el punto de vista lógico no necesitan ser exploradas por él, gracias a restricciones innatas que guían el aprendizaje y restringen opciones que de otro modo deberían ser consideradas; en otras palabras, tales restricciones le simplifican la tarea.

La resolución de este problema de aprendizaje en tiempo finito (siguiendo a Gold 1967) depende de la cantidad de mecanismos usados por la teoría para dar cuenta de ese objetivo, entendido como el trazado desde el estado inicial a los estados finales, y de la restricción efectuada sobre ese número de gramáticas ilimitadas accesibles al alcance del niño según esos mecanismos. De esta manera, un número alto de mecanismos particulares ofrecidos o permitidos por la teoría general se proyectará de una manera numéricamente muy alta, produciendo muchas opciones, con lo que será más difícil dar cuenta de la adquisición dadas las hipótesis compatibles con los datos hasta alcanzar la convergencia.

Ésa era precisamente la situación producida en las primeras versiones de la GG, las cuales usaron una gran cantidad de mecanismos específicos y particulares de lenguas para obtener adecuación descriptiva, pero en la misma medida no se pudo acceder a la explicativa, ese estadio inicial común representado por medio de la GU (cfr. Longa 1999a). En otras palabras, la GG no pudo llevar a unos cauces razonables la factibilidad para el aprendizaje en tiempo finito, porque el elevado número y clase de reglas usadas (especialmente transformacionales y de reescritura) originaba muchas hipótesis accesibles, con lo que la especificación de la teoría de la GU estaba carente de contenido.

Sin embargo, P&P cambió radicalmente la situación, puesto que uno de sus rasgos principales consistió en la escasa cantidad y gran generalización de mecanismos usados, culminando la eliminación de la mayor parte del aparato expresivo comenzada a fines de los años 60. De este modo, se pudo reducir hasta límites aceptables el número y la clase de gramáticas accesibles proyectadas por el marco general, al

² A veces se malinterpreta el marco innato, ejemplo de lo cual es Lieberman (1998: 128), que señala que "The power of cognitive processes like associative learning and imitation is routinely ignored in many theoretical discussions of language acquisition". Sin embargo, la existencia de un problema lógico no implica que en todas las áreas haya este problema; la GG no niega esos procesos, sino que afirma que no son suficientes, como escriben Lightfoot (1999: 63): "Nobody denies that the child must extract information from her environment (...). My point is that there is more to language acquisition than this" o Pinker (1997: 32): "If the mind has a complex innate structure, that does not mean that learning is unimportant". Cfr. igualmente Cowie (1999: 40), referencia que constituye una crítica del innatismo aunque se reconoce lo señalado.

ser éste mucho más restringido, lo cual, por razones evidentes, pudo limitar las posibilidades ilimitadas desde la perspectiva lógica, insertando el problema de aprendizaje en cauces factibles. En otras palabras, se pudo formular una teoría de la GU con verdadero contenido (cfr. Longa 1999a).

En la nueva perspectiva, que básicamente implica el paso de reglas a principios, existen diferentes principios universales, que remiten a aspectos innatos³, con un menú de opciones o parámetros cada uno de ellos, usualmente dos, también dependientes del estado inicial. Esa doble opción para cada principio va siendo fijada en cada caso conforme el niño accede a la experiencia específica relevante (exposición a las diferentes lenguas), resultando al final del proceso, tras la fijación de todos los parámetros, el estado final correspondiente. La labor del niño consiste en fijar esas posibles opciones de acuerdo con los disparadores (o *triggers*; cfr. Lightfoot 1991, Gibson & Wexler 1994) que constituyen tales datos, bien abiertos (Hyams 1986), o bien en forma de señales más abstractas (sobre esta última visión, cfr. en diferentes vías Morgan 1986, Lightfoot 1999). Cada combinación diferente de parámetros constituye una gramática posible en este marco, aunque debido a la restricción mencionada, el número potencialmente alto de ellas no es un problema de aprendizaje, puesto que, como escribe Freidin (1994: 680), "linguistic theory allows only a severely limited number of grammars - perhaps only one for each given language". Este modelo, además de la resolución del problema lógico, está mucho más en consonancia con el carácter del crecimiento del lenguaje señalado en ocasiones por Chomsky: no es algo que el niño hace, sino algo que le pasa. El anterior sistema de reglas presuponia una naturaleza mucho más activa del niño (ensayar hipótesis de entre un número muy alto), y sobre todo, mucho menos factible.

Existen diferentes aproximaciones a la fijación de esos parámetros, si bien todas coinciden en el planteamiento general: por ejemplo, el algoritmo de aprendizaje de disparadores (Gibson & Wexler 1994; cfr. Fodor 1998 donde se modifica ese sistema), el aprendizaje basado en señales (Morgan 1986, Dresher & Kaye 1990, Dresher 1999, Lightfoot 1999), algoritmo genético (Clark 1992, Clark & Roberts 1993) o el algoritmo de paso al azar, formulado en términos de cadenas de Markov, de Berwick & Niyogi (1996). Exponiendo someramente uno de los más detallados, el modelo de Gibson & Wexler, estos autores asumen que los disparadores que son la base para fijar los parámetros informan sólo sobre un único parámetro en cada caso, y no sobre varios, aspecto tomado de Clark (1992). El algoritmo mencionado, *Triggering Learning Algorithm* (p. 409) pretende ser aplicado a cualquier aspecto de la gramática (frente por ejemplo a Dresher & Kaye 1990, que formulan su aproximación para secciones bien definidas, y que por ello, no es válida para cualquier área). El algoritmo consiste en una búsqueda al azar de parámetros, y guiada por error; esto es, en respuesta a un error, entendido como la imposibilidad de analizar una secuencia con la gramática postulada por el niño (una secuencia es analizable si puede ser generada por la gramática adoptada por el niño en ese momento), se selecciona un paráme-

³ En una doble posibilidad, como señala Lightfoot (1999: 60): bien insertados directamente en el genoma, o bien en el sentido de propiedades de desarrollo de los organismos, o reglas epigenéticas (en Wilson 1999 se expone una panorámica de esta segunda posibilidad en relación a muy variados dominios).

tro al azar, variando su valor. La formulación es como sigue (Gibson & Wexler 1994: 409),

The Triggering Learning Algorithm (TLA)

Given an initial set of values for n binary-valued parameters, the learner attempts to syntactically analyze an incoming sentence S . If S can be successfully analyzed, then the learner's hypothesis regarding the target grammar is left unchanged. If, however, the learner cannot analyze S , then the learner uniformly selects a parameter P (...), changes the value associated with P , and tries to reprocess S using the new parameter value. If analysis is now possible, then the parameter value change is adopted. Otherwise, the original parameter value is retained.

El algoritmo, pues, depende de la existencia de disparadores, los datos que indican el valor en que debe ser fijado un parámetro. El éxito se obtiene para un principio si por azar se selecciona un parámetro que permite que la gramática del niño pueda incorporar ese disparador, generándolo, y eliminando así el desfase producido por la imposibilidad de generar esa secuencia con la gramática que en ese momento dispone del niño, ya que según Gibson & Wexler (1994: 410), "the learner has no knowledge of which parameters are relevant to a given input".

Independientemente de uno u otro de los enfoques señalados, el problema lógico de la adquisición lingüística se ha encauzado definitivamente por medio del modelo de P&P, puesto que la respuesta a él es positiva: frente a modelos anteriores, en éste la gramática correcta es aprendible, siendo eliminadas muchas posibles por medio de restricciones sustantivas. Nótese que si efectivamente hay pobreza de datos, la predicción consiste en que precisamente muchas posibilidades sean eliminadas gracias a la carga innata, lo que hace la tarea mucho más asequible y fácil de lo que sería en otras circunstancias.

3. LA TEORÍA DE LA OPTIMALIDAD

Esta reciente teoría generativa surgió (Prince & Smolensky 1993) como un intento de eliminar farragosas derivaciones fonológicas, donde reglas y restricciones específicas debían ser aplicadas en diferentes niveles. Sin embargo, poco después comenzó a ser postulada en morfología y sintaxis, siendo según sus seguidores una alternativa real al modelo de P&P. De manera todavía más reciente, ha comenzado igualmente a ser aplicada al aprendizaje formal (cfr. *infra*).

Brevemente caracterizada, esta teoría sostiene los siguientes asertos centrales: (I) existencia de un conjunto universal de restricciones, (II) ordenación de esas mismas restricciones en diferentes jerarquías o *rankings* en las lenguas (de las que resulta la variación interlingüística), (III) posibilidad de que los *outputs* o resultados producidos por la jerarquía sean gramaticales a pesar de que se viole alguna restricción (colocada en los lugares más bajos de esa jerarquía), lo cual se produce para garantizar el cumplimiento de otra restricción colocada en los lugares superiores, y (IV) separación de lo gramatical y lo agramatical en un solo paso, efectuada por la jerarquía de restricciones. Por tanto, la idea central de Optimalidad consiste en que cada lengua

coloca las restricciones universales en un orden particular; de este conflicto entre la ordenación de restricciones resulta la variación.

El formato de TO es como sigue: se parte de un componente denominado INPUT, cuya única restricción está constituida por el que los objetos conformados en él deben atenerse al vocabulario (categorías, etc.) ofrecidas por la GU (esto es, deben ser representaciones lingüísticas). Los *inputs* entran en el componente GENERADOR, que se encarga de crear *outputs* potenciales a partir del *input* dado. Este componente es bastante creativo, por lo que se pueden añadir o suprimir elementos del *input*. Una vez creado el conjunto de resultados potenciales para el *input*, estos son juzgados en el componente EVALUACIÓN por medio del conjunto de restricciones universales CONJUNTO, de manera que de esa evaluación se deriva el resultado óptimo para ese *input*, que es el candidato ganador. Éste es precisamente el que satisface un mayor número de restricciones (bien todas, o bien las situadas en los lugares más altos de la jerarquía específica de esa lengua). Por ello, debe resultar un ganador, pero sólo uno (en casos de opcionalidad deben existir dos diferentes jerarquías implicadas). Otra característica importante estriba en que no existe ningún tipo de derivación, puesto que no se usan reglas, sino que se parte de sus resultados directamente, y tampoco existe ninguna representación intermedia entre el *input* o representación inicial y los *outputs* o representaciones finales (lo cual está motivado por el deseo de efectuar la separación entre lo gramatical y lo agramatical en un solo paso).

Tales son, de manera muy breve, los supuestos básicos de esta teoría (para mayores detalles, cfr. los diferentes trabajos en Archangeli & Langendoen (eds.) 1997, Barbosa et al. (eds.) 1998, Gilbers & De Hoop (eds.) 1998, y Longa 1999b, en prensa).

4. LA TEORÍA DE LA ADQUISICIÓN Y EL PROBLEMA LÓGICO EN OPTIMALIDAD

Como señalé antes, la cuestión del problema lógico estriba en determinar la aportación del estado inicial al proceso de adquisición lingüística. Frente a teorías empiristas (o de otro tipo) que defienden la suficiencia de los datos como para poder aprender una lengua directamente a partir de ellos, no concediendo así ningún papel a una posible aportación del estado inicial (cfr. Sampson 1997), TO defiende claramente la insuficiencia de los datos, y por ello, la existencia de un problema lógico. En palabras de Pulleyblank & Turkel (1998: 400),

One of the central problems for a theory of learnability is that the evidence available to the language learner is too sparse to account for the complexity of the adult speaker's knowledge of language (...); in Optimality Theory, such adult knowledge is achieved through the establishment of the correct ranking of universal constraints. The problem for the learner (...) is to arrive at the required grammar on the basis of a potentially deficient set of positive data from the language being acquired.

Este problema lógico es considerado desde la misma perspectiva que en P&P, y en general, en GG: por medio de una carga innata que permite suplir esas deficiencias. Por

tanto, también esta teoría pone en lugar central una GU (denominada CONJUNTO, el conjunto de las restricciones universales) decisiva para el proceso de adquisición, aunque dadas las características de Optimalidad antes enunciadas, existen divergencias apreciables entre los dos enfoques generativos, derivadas de los respectivos rasgos generales de formato: mientras en P&P la GU, o componente innato, está modelizada mediante una serie de principios universales parametrizados, y lo aprendido consiste en el léxico junto a la fijación de los principios en una de las dos posiciones, en Optimalidad se postula un trazado entre el estado inicial y el estable en el que la carga innata especifica la función GENERADOR, y además unas restricciones universales (que pueden ser infringidas). Lo que debe aprenderse es el lexicón que ofrece las formas subyacentes en los *inputs* de esa lengua, a partir de los que se generan los *outputs*, y el orden de las restricciones en jerarquías determinadas (Tesar 1998a: 421). Sin embargo, a pesar de estas diferencias entre ambos marcos, los dos comparten los rasgos principales (especialmente, el reconocimiento de la aportación innata), por lo cual Pulleyblank & Turkel (1998: 399) afirman que la diferencia con P&P es de grado, no de clase.

Lo señalado implica que según Optimalidad, el aprendiz, a partir de los datos que recibe y que debe analizar, debe ir confeccionando la jerarquía de restricciones propia de esa lengua, o del fenómeno al que alude el dato recibido. Si el niño no tiene forma de analizar una secuencia según la jerarquía (esto es, la gramática) hipotetizada en ese momento, esa jerarquía debe ser modificada hasta que sea capaz de analizar esa secuencia; el final del proceso llega en el momento en el que hay convergencia plena entre la gramática del estado final y la gramática del niño.

A pesar de que los aspectos esbozados constituyen las líneas generales de la visión formal de la adquisición según este marco, existen, al igual que en P&P, diferentes enfoques que, teniendo coincidencias centrales (la jerarquía, su reformulación, etc.), tratan el trazado entre ambos estados de manera algo diferente. Los dos principales en Optimalidad son los representados por un lado por Pulleyblank & Turkel (1998) y por otro, por Smolensky & Tesar (Smolensky 1996, Tesar 1998a,b, Tesar & Smolensky 1998). Los primeros optan por un modelo de aprendizaje representado por un algoritmo genético (así denominado por estar basado en analogías con la recombinación biológica producida por deriva genética; cfr. Forrest 1993 para una perspectiva general y, más en concreto, Turkel 1996), cuyo carácter, en todo caso, no concierne específicamente a Optimalidad, puesto que fue aplicado previamente a la fijación de parámetros en Clark (1992).

Este algoritmo procede, a grandes rasgos, de la siguiente manera: un conjunto de jerarquías es generado al azar, constituyendo la población inicial. El niño selecciona pares *input/output* para cada dato al que accede, debiendo ser el *output* consistente con ese dato al alcance del aprendiz. Estos pares permiten apreciar la eficacia de cada gramática (esto es, de cada jerarquía de restricciones) de la población: cuanto más bajas sean en jerarquía las restricciones violadas, más ajustada es la gramática, y por ello, más cerca está esa gramática con la gramática del estado estable.

Con respecto al segundo enfoque, representado por Tesar & Smolensky, me detendré con mayor profundidad en su exposición, por ser más completo y elaborado,

teniendo en cuenta además que de él se intenta derivar también el desfase existente entre producción y recepción en el niño (cfr. Smolensky 1996 específicamente sobre este aspecto⁴; como es bien sabido, en la adquisición el niño es capaz de entender bastante más de lo que en ese momento puede producir). Frente al modelo anterior, algoritmo genético, este procedimiento no funciona por búsqueda al azar, sino que es dirigido por error, como veremos posteriormente. Otro factor importante consiste en que en esta perspectiva se insiste en la necesidad de no tratar el proceso de adquisición (relación entre formas abiertas o datos accesibles y gramáticas) de manera única, como usualmente se hace, sino teniendo en cuenta la mediación de las descripciones estructurales de los datos hipotetizadas por el niño entre las formas abiertas y las gramáticas (Tesar 1998b: 132, Tesar & Smolensky 1998: 236); así, el problema del aprendizaje se descompone en esta perspectiva en dos factores: por un lado, la determinación de la relación entre las descripciones estructurales plenas y las gramáticas y, por otro, la falta de información de las formas abiertas como para determinar sus descripciones estructurales plenas. Como veremos, se defiende un procedimiento de aprendizaje iterativo, donde se va desde la utilización de una jerarquía para estimar la estructura oculta posible de un dato al que accede el aprendiz (que apunta hacia el candidato ganador, ya que se asume la no maliciosidad de los ejemplos; Tesar & Smolensky 1998: 246) hacia la utilización de esa estructura oculta posible para reordenar la jerarquía, y viceversa, proceso en el que la clave es el procedimiento de degradación de restricciones (cfr. *infra*).

Como señalaba, se postulan dos diferentes procesos: análisis dirigido a la producción y análisis interpretativo (cfr. nota 4): el primero traza un *input*, o forma subyacente hacia su descripción estructural óptima, dada una gramática hipotetizada, mientras que el segundo analiza una estructura abierta (dato recepcionado) en una descripción estructural y estructura oculta, dada igualmente una gramática. Por ello, aprender una gramática es determinar la jerarquía de restricciones a partir de las descripciones estructurales, razón por la que éstas tienen un papel mediador central en el proceso: nótese a este respecto que ambos procesos consideran una descripción estructural (o una parte), aunque en cada caso cambia la forma a la que alude esa descripción: bien forma subyacente, bien forma abierta. El niño empieza con una jerarquía inicial, usándola para asignar descripciones estructurales a las formas abiertas que recibe; esas definiciones se emplean para cambiar la jerarquía en caso de desfase, y la nueva resultante sirve otra vez para determinar una nueva estructura oculta: de ahí el carácter iterativo del proceso, que finaliza exitosamente cuando esas iteraciones

⁴ Precisamente, la posibilidad de que TO dé cuenta de este desfase es presentado por Smolensky (1996: 730) como evidencia a favor de la superioridad de este modelo de adquisición sobre el de P&P: en este último, a su juicio es difícil apreciar por qué con una sola gramática el niño puede usar un conjunto reducido de elecciones de parámetros en su producción y, sin embargo, es capaz de procesar bien formas adultas que requieren varios conjuntos. Si por el contrario la gramática se concibe como una jerarquía violable, el niño procesará bien lo que aún no está en disposición de producir, porque las diferencias entre los conjuntos competidores permiten interacciones ocultas al niño en la producción, pero decisivas en la comprensión. Esto es, Smolensky rechaza que haya que asumir la existencia de dos gramáticas diferentes al tiempo (para producción y recepción), sino que presupone una sola, aunque con posibilidades diferentes derivadas de diferentes ordenaciones, cual es el espíritu de TO.

convergen, siendo la jerarquía del todo consistente con las formas abiertas o superficiales. De ahí el papel central de las descripciones para aprender la gramática (esto es, el orden de restricciones), y también central con respecto a los dos procesos señalados.

Tras estas consideraciones, hay que incidir en cómo se va produciendo la conformación y reformulación de las jerarquías para obtener la convergencia plena. El sistema comienza a trabajar con una serie de restricciones postuladas sin ordenación alguna, esto es, sin estar dispuestas con arreglo a ninguna jerarquía, sino indiferenciadas en un único estrato. La labor del aprendiz consiste en ir formulando jerarquías de las restricciones sobre la base de los datos accesibles, distribuyendo éstas, por tanto, en estratos diferentes, de manera que sea posible comparar candidatos alternativos con arreglo a esas jerarquías hipotetizadas. Esta visión, según la cual el niño formula diferentes jerarquías de las restricciones, implica que es necesaria una labor de re/ordenación de ellas.

Para esta labor se ha postulado un algoritmo denominado 'degradación de restricciones' (*Constraint Demotion*; cfr. Tesar 1998a: 424 y ss., 1998b: 138, Tesar & Smolensky 1998: 238 y ss.), que trabaja guiado por error⁵, y que garantiza la deriva hacia la obtención de la jerarquía correcta. Si se produce un error, actúa el algoritmo señalado, cuya formulación es como sigue (Tesar & Smolensky 1998: 240),

Constraint Demotion

For any constraint C assessing an uncanceled winner mark, if C is not dominated by a constraint assessing an uncanceled loser mark, demote C to immediately below the highest-ranked constraint assessing an uncanceled loser mark.

La idea de este algoritmo es bien simple: pretende asegurar que cada *output* ganador de entre el conjunto de candidatos sea más armónico que el resto de los competidores de ese conjunto, donde orden armónico significa simplemente el orden adecuado de restricciones (Tesar & Smolensky 1998: 233); de este modo, el candidato óptimo y, por ello, ganador, es aquél que presenta un orden más armónico de restricciones. Así, el algoritmo implica que, al existir un desfase entre la gramática que hipotetiza el aprendiz y el análisis de una forma, el orden debe ser modificado de manera que el nuevo corrija esa diferencia, hasta eliminarla. En relación con esto, recuérdese que el candidato ganador no tiene por qué satisfacer todas las restricciones, sino que puede violar alguna, con tal de que esté colocada en los lugares más bajos de la jerarquía. Esto significa que las violaciones de restricciones por parte del candidato ganador deben quedar por debajo de las violaciones de restricciones acaecidas en los perdedores. Por tanto, la restricción R infringida por el ganador debe ser descendida a su lugar adecuado en la jerarquía, procedimiento reiterado en cada par de restricciones hasta fijar el orden⁶; cuando hay desfase entre la forma abierta y su análisis subyacente con

⁵ Por error hay que entender, por ejemplo, que la interpretación de una forma abierta obtenida mediante un análisis interpretativo no se empareje a la perfección con la descripción estructural de la forma subyacente generada gracias al análisis dirigido a la producción.

⁶ Según Tesar & Smolensky (1998: 244-5) sólo un descenso de restricciones, y no un ascenso, puede garantizar la convergencia con el orden adecuado. La razón consiste en que si se postulara una promoción de restricciones, no se conocería cuál de ellas debe ser elevada; sin embargo, si el

respecto al orden hipotetizado por el aprendiz, se baja una restricción: si éste recibe un ejemplo E, computa su propio análisis E' según su jerarquía en ese momento. Si E' difiere de E, hay posibilidad de aprendizaje, que pasa por aplicar la degradación de restricción; si E y E' coinciden, no hay posibilidad ni necesidad de él.

Este algoritmo permite resolver, según sus proponentes, un problema evidente causado por la importancia decisiva de las definiciones estructurales: cómo accede el niño a su formulación sin haber confeccionado aún la jerarquía de restricciones del estado final; en palabras de Tesar (1998b: 132), indefinición de las definiciones estructurales por parte de las formas abiertas. Sin embargo, este autor defiende que el algoritmo de degradación puede explicar ese aspecto: el niño estima la estructura oculta de una forma recibida, de lo cual resulta una descripción estructural como hipótesis, que es usada para modificar la jerarquía aplicando una o varias degradaciones. De nuevo, pues, se aprecia la importancia de la iteración en el proceso, ya que por un lado una jerarquía sirve para estimar una estructura oculta y por otro, esa estructura oculta se emplea para reordenar la jerarquía.

En (1998: 242 y ss.), Tesar & Smolensky ofrecen un ejemplo de degradación de restricciones referido a una estructura silábica artificial, mostrando que ese procedimiento garantiza la obtención de la jerarquía correcta con independencia del orden de restricciones con el que se empiece (de ahí que sólo el comienzo sea al azar, y lo demás, por error). El procedimiento de degradación asegura una convergencia rápida, fundamental para todo modelo que trate de satisfacer la condición de factibilidad de aprendizaje. Teniendo en cuenta las diferentes posibilidades de ordenación de restricciones, el número de jerarquías posibles, enumeradas por fuerza bruta, consiste en una función factorial del número de restricciones planteadas. Esto es, en un sistema con 5 restricciones existen 120 posibilidades diferentes de ordenación, mientras que con 10, el número sube hasta 3.628.800. Incrementado el número en una más, 11, tenemos 39.916.800 ordenaciones diferentes. Sin embargo, en una simulación relacionada con el acento métrico, en Tesar (1998b: 140 y ss.) se muestra que el algoritmo llega a un orden correcto tras solamente aplicarse 10 veces la degradación. De hecho, Tesar (1998a: 431) señala que el número más alto de errores necesitados para poder obtener la convergencia plena con la jerarquía correcta, dado un número N de restricciones, es $\frac{1}{2} (N_{restr} - 1) N_{restr}$. Siguiendo con el ejemplo anterior de 11 restricciones, de entre los casi 40 millones de posibilidades, el procedimiento de degradación guiada por error asegura que la ordenación correcta se encontrará tras haber considerado como máximo 55 errores o disparidades entre datos y análisis óptimo con respecto a la gramática hipotetizada en ese momento por el niño, lo que evita analizar todas las posibilidades una a una hasta encontrar al azar la adecuada.

Otro rasgo importante consiste en el tipo de experiencia que asume este marco: la única evidencia disponible que tiene el aprendiz es la positiva. Sin embargo, frente a P&P, en TO se afirma que la evidencia negativa es ofrecida por el propio niño, me-

procedimiento es de descenso, el problema se soluciona bajando la restricción o restricciones violadas por el candidato ganador. De esta manera, el aprendizaje es guiado por error, y no al azar, y sobre todo, tampoco enumerando por fuerza bruta todos los órdenes posibles, lo cual sería el caso si hubiera promoción de restricciones.

diante el análisis de cada conjunto candidato, del que sale un ganador y varios perdedores⁷. De hecho, el propio diseño del procedimiento es central para esta asunción: el aprendiz no puede reordenar restricciones mediante la degradación de las correspondientes con el fin de alcanzar el orden correcto si no es comparando los candidatos óptimos de los conjuntos de candidatos generados a partir de los datos positivos con el resto de los competidores, que se erigen así en evidencia negativa implícita.

Para resumir, pues, el procedimiento es iterativo e interactivo, al estar posibilitada la adquisición por la interacción entre producción y recepción: mientras en la primera se computa la descripción estructural óptima para una forma subyacente, en la comprensión sucede al revés, al ser seleccionada la descripción estructural de una forma abierta que sea óptima con respecto a la jerarquía. Si la interpretación de una forma abierta no equivale a la descripción estructural de la forma subyacente generada por medio del análisis hacia la producción, se produce un error, de modo que se provoca una alteración de la jerarquía asumiendo que el candidato ganador es el análisis interpretativo y el perdedor es el dirigido a la producción (Tesar 1998b: 140). De ahí que los errores se puedan corregir mediante la comprensión (Smolensky 1996: 729). Cuando finalmente no hay ningún desfase, se acaba el proceso de adquisición, entendido como adquisición de la gramática dado un léxico, y del léxico dada una gramática.

5. ALGUNOS COMENTARIOS CRÍTICOS SOBRE EL APRENDIZAJE EN LA PERSPECTIVA DE TO

La idea global en la que se basa esta teoría, la gramática como orden o jerarquía particular de restricciones universales, no se puede dejar de reconocer como interesante. Sin embargo, existen algunos aspectos problemáticos derivados de la visión manejada por esta corriente. Señalaré algunos de ellos en relación al enfoque que más extensamente se ha presentado (si bien la mayor parte de estos aspectos son aplicables al otro enfoque).

5.1. Viabilidad global y parcial de la teoría lingüística

Como se apuntó en el apartado 3, Optimalidad es una teoría realmente reciente. Si tenemos en cuenta que su surgimiento tuvo que ver inicialmente con la fonología, se podrá apreciar que su proyección en sintaxis y en adquisición es aún más novedosa. En sintaxis, los trabajos inscritos en el marco se dedican a comparar un mismo fenómeno en diferentes lenguas, pero el aspecto central desde esta perspectiva consiste en la elaboración de una gramática global; dicho de otro modo, los trabajos inscritos en el marco de TO aluden a jerarquías en relación a los fenómenos estudiados, pero pre-

⁷ En P&P se acepta en ocasiones, al lado de la evidencia positiva, una evidencia negativa implícita, motivada por la ausencia de una determinada forma en ciertos contextos, factor que, frente a Optimalidad, no tiene que ver con ningún tipo de análisis subyacente del niño. Sin embargo, debe destacarse que no hay consenso al respecto, puesto que si bien autores como Chomsky (1981) o Lasnik (1989) alegan ese tipo de evidencia, otros, como especialmente Lightfoot (1999), señalan que no está bien motivada.

cisamente es en una gramática global donde se podrá apreciar si el modelo realmente funciona, debido al descubrimiento de si se mantiene en pie cuando se traten las interacciones entre las diferentes jerarquías de que se compone una gramática. A este respecto, P&P ha mostrado sólidamente que puede formular una gramática global con garantías, soportando el modelo las interacciones entre los diferentes componentes (principios, parámetros, etc.).

Esto mismo se puede aplicar a la adquisición: los casos mediante los que se ejemplifica el procedimiento de adquisición se dedican a cuestiones de estructura silábica, tanto artificial como natural, y a poco más (en sintaxis apenas hay tratamientos). Se asume que el niño aprende un segmento, pero en relación con lo señalado antes, el niño aprende al tiempo, conforme recibe e interioriza los datos de diferentes segmentos. En este sentido, la teoría de la adquisición va a remolque (como en cualquier teoría formal) de la propia teoría lingüística; por ello, hasta que no se muestre la viabilidad de la teoría, en la misma medida no será posible mostrarlo en relación a la adquisición. Como escribe Newson (1998: 315), "Optimality Theory is an infant and we do not know what it is going to be when it grows up".

Pero además de que actualmente TO está aún lejos de poder demostrar un tratamiento global de la sintaxis, hay algún caso específico que no parece ser cubierto por las predicciones de la teoría, por ejemplo referido a los infinitivos conjugados (para detalles, cfr. Longa en prensa) y a su vinculación con una teoría de la marca. Las formas verbales y su concordancia suelen ocurrir en las lenguas en formas finitas, y muy raramente en las infinitas (Moravcsik 1988: 95). En términos de Chomsky (1981: 52), la asociación por defecto en las formas finitas consiste en el valor positivo del elemento T(tiempo) con la especificación de C(concordancia), por tanto, [+T, +C]. Inversamente, la otra opción defectiva es el valor negativo en ambos casos, situación común en infinitivos. Sin embargo, no todas las lenguas presentan estas opciones no marcadas: por ejemplo, [+T, -C] se produce en chino, y desde la perspectiva contraria, otra situación marcada está conformada por [-T, +C] en infinitivos de lenguas como gallego o húngaro, rasgos de concordancia asociados a un elemento Tiempo negativo (cfr. Raposo 1987 y Longa 1994 para detalles).

Veamos qué predicciones puede ofrecer Optimalidad. En principio, sería esperable que las diferencias entre infinitivos conjugados y no conjugados se pudieran explicar mediante dos órdenes diferentes, que caracterizaran esta oposición tanto dentro del gallego (infinitivos con y sin morfema de concordancia) como entre gallego y, por ejemplo, castellano. Las restricciones relevantes para este fenómeno serían las siguientes: una que alude al movimiento más corto, versión en TO de la economía chomskyana, que es violada, pues V debe subir de FV a Flex y después a C. Esta violación se debe al cumplimiento de la restricción 'Satisfacción', mediante la que un elemento debe moverse para comprobar sus rasgos. Ambas restricciones están en conflicto, lo que va bien con el espíritu de la teoría. Sin embargo, si atendemos a más principios, no parecen existir dos órdenes diferentes: por ejemplo, la restricción 'Flex' alude a la caracterización normal de los infinitivos como [-T, -C]. Esta restricción interactúa con la vertiente en Optimalidad del filtro casual según la que una FN realizada debe tener caso, llamada 'Caso'. La restricción 'Flex' es respetada en los infini-

tivos sin marca de concordancia, pero no en los que las presentan, porque la forma tiene marcas personales.

Recuérdese que Optimalidad tiene como aserto básico que cuando una restricción es quebrantada, esta infracción está encaminada a la satisfacción de una restricción colocada en un lugar más alto de la jerarquía. Por ello, lo esperable sería que 'Flex' se violara para obtener un resultado diferente del producido en los infinitivos sin marca personal, relacionado con la posibilidad de tener un sujeto léxico. Sin embargo, esto no es así: se viole o no esa restricción, el infinitivo puede asignar caso: recordemos que los infinitivos no conjugados pueden asignar caso a un sujeto mediante el movimiento de V a un operador (Rigau 1995) que junto a P (o a otro legitimador) ofrece las condiciones necesarias para la asignación (lo mismo sucede en húngaro, donde también existe esta opcionalidad), con lo que hay dos ordenaciones iguales para dos fenómenos diferentes, y no dos órdenes diferentes, pero esa única ordenación tiene, sin embargo, diferencias en infracciones (lo mismo se podría destacar con respecto a una restricción general, 'Fidelidad' (*faithfulness*), respetada en un caso pero no en otro, dependiendo del *input* del que se parta, con o sin marca de concordancia). Por ello, una misma jerarquía puede dar dos candidatos óptimos, aunque hay diferencias de infracción en al menos dos restricciones.

Si aplicamos al terreno de la adquisición este resultado, aceptando su corrección, la inexistencia de dos jerarquías para dos diferentes fenómenos (infinitivos con y sin concordancia) y la presencia de dos candidatos óptimos en una única jerarquía, aunque con diferencias de infracción en dos restricciones, implica que, además de no casar con los presupuestos de la teoría lingüística de Optimalidad, no hay procedimiento de aprendizaje imaginable para los infinitivos conjugados, lo cual es algo indeseable, sobre todo porque en este caso ni siquiera es necesario invocar evidencia negativa, siendo relevantes sólo los datos presentados, esto es, la evidencia positiva. De hecho, en P&P el tratamiento de este fenómeno en cuanto al aprendizaje es bien sencillo: simple extensión marcada de una opción por defecto, o no marcada. Cuando el niño tiene datos al respecto (secuencias recibidas con infinitivos que disponen de concordancia), que conforman el disparador adecuado, abandona esa opción no marcada. Sin embargo, en Optimalidad no hay procedimiento en el caso discutido: la adquisición es concebida como análisis de *inputs* y *outputs*, y de descripciones estructurales, que tienen un papel mediador central, como vimos, con respecto a la confección de jerarquías. En este caso, no se podría aplicar la degradación de restricciones con éxito, puesto que valdría para cualquiera de las dos posibilidades un orden, aunque habría diferencias de violación en algunas restricciones.

Por ello, mencionaba antes que en primer lugar TO deberá revelarse como viable en el terreno de la propia teoría lingüística; en caso contrario, será inviable su tratamiento de la adquisición.

5.2. La generación de conjuntos candidatos

Uno de los componentes del modelo es la función GENERADOR, que se encarga de ofrecer un conjunto de candidatos a partir de un *input* (o de un dato abierto reci-

bido por el aprendiz). Esta función se asume como universal, no aprendida. En la universalidad del equivalente de esta función (el sistema computacional) también coincide P&P. Sin embargo, en este último modelo lo que hay que aprender son los rasgos de las piezas léxicas, para que sean aplicados a ese sistema computacional, mientras que en Optimalidad es necesario conformar diferentes candidatos, que son la base para el verdadero objetivo del modelo, la comparación y valoración en la jerarquía. De ahí que este modelo semeje más recargado y complejo.

Es importante notar, además, dos cuestiones: la primera tiene que ver con la irrestricción del componente GENERADOR: éste puede generar infinitos candidatos a partir de un *input* (conforme se quebrante el requisito de fidelidad en mayor o menor grado). Esta sobregeneración, que no es relevante para la propia teoría lingüística, lo es, sin embargo, en cuestiones de adquisición, porque causa que el problema sea mucho más complejo en cuanto a la factibilidad computacional: no porque el aprendiz tenga que valorar todas las posibles generaciones en el conjunto candidato, sino porque la valoración de unas pocas implica una decisión de escogerlas sobre el resto, para lo cual el único criterio parece ser la similitud con la forma correcta. En otras palabras, no es diáfana la clase de los competidores que deben entrar en la evaluación. El segundo problema tiene que ver con la nula atención a los mecanismos por los que se produce la generación: estos no se consideran, abordando únicamente la selección del *output* óptimo a partir de un conjunto candidato conformado de golpe. En este sentido, se deja de lado algo muy importante, la función combinatoria en términos de Uriagereka (1998: 168-9), para la que una teoría de la adquisición debe ofrecer una explicación. Aunque esta teoría niega cualquier tipo de derivación (cfr. Archangeli 1997), a la luz de lo señalado, parece evidente que en el componente que genera es central un procedimiento de derivación, con respecto al que, sin embargo, no se explicita nada, excepto que puede ser infinito.

Cada conjunto candidato debe ser evaluado para escoger un ganador, pero, además de lo señalado en el párrafo anterior, esta hipótesis recarga mucho el modelo, al implicar no ya una valoración para establecer cada jerarquía, sino para cada aplicación de la degradación de restricciones. Al tiempo, debe mencionarse un aspecto relativo al orden inicial de restricciones manejado: aunque los autores afirman que es irrelevante por dónde se empieza, pues la degradación consigue converger hacia la jerarquía correcta, sin embargo no sólo hay una jerarquía implicada en cada gramática, sino diferentes para diferentes fenómenos. Esto significa que, además de tener que determinar jerarquías a la vez, el niño debe ser consciente de que diferentes jerarquías tienen que ver con diferentes restricciones. Por el contrario, en P&P el niño simplemente espera unos disparadores (o señales más abstractas) que vayan fijando los principios en posiciones determinadas.

5.3. Espacio de gramáticas posibles

Este aspecto es central para cualquier teoría que desee ofrecer un modelo factible de aprendizaje. Por señalarlo en pocas palabras, los mecanismos postulados conforman un número de hipótesis accesibles, o gramáticas posibles. El ejemplo clásico podría ser el orden de las transformaciones en los primeros estadios de la GG: postu-

lando 10 transformaciones, y asumiendo que deben estar ordenadas extrínsecamente, si no se ofrecen mayores especificaciones, esto significa que el espacio de hipótesis accesibles consiste en una función factorial del número de transformaciones: en este caso, 3.628.800 posibles maneras de ordenar, que el niño supuestamente debería explorar (nótese que esta hipótesis trabajaba sobre pares estrictamente adyacentes de reglas).

En un sistema donde existen varias restricciones, también a priori habría muchas posibilidades, lo cual, como en el ejemplo de las transformaciones, sólo dejaría la posibilidad de enumeración por fuerza bruta, buscando al azar hasta lograr encontrar la jerarquía adecuada. De ahí la proposición de un algoritmo de degradación de restricciones guiado por error, que evita la búsqueda al azar, o por fuerza bruta, enumerando todas las posibilidades. Por esto, según los defensores de esta óptica, este modelo de aprendizaje es superior al postulado en P&P. Por ejemplo, Tesar & Smolensky (1998: 229) señalan que un número tan corto como 20 parámetros posibilita una cifra de un millón de gramáticas posibles, obteniendo la correcta mediante el sistema de que, en respuesta a un error, se varía el valor de un parámetro seleccionado de manera azarosa. Sin embargo, aquí surgen diferentes aspectos. En primer lugar, en el modelo de P&P, las diferentes posibilidades de combinación no constituyen un problema de aprendizaje con consecuencias semejantes a las de modelos previos. Este modelo ha sido restringido como para que esas opciones simplemente marquen un espacio de hipótesis, pero una gramática es aprendible en esta perspectiva, ya que hay un resultado de adquisición estrictamente finito. Esto es, el número de posibilidades no es un problema de aprendizaje verdadero, sino más bien las interrelaciones entre los parámetros. En Gibson & Wexler (1994: 407-9) se expone que el espacio de gramáticas posibles se simplifica mucho mediante la disposición paramétrica.

En Optimalidad, la situación es bien diferente, ya que no puede existir una única jerarquía de restricciones a partir de la que se dé cuenta de la adquisición de toda la sintaxis de una lengua (cfr. 5.1. y 5.2.). Por ello, incluso suponiendo que el modelo pudiera efectivamente caracterizar la adquisición, el niño necesitará conformar bastantes jerarquías diferentes. De ahí que el número de posibilidades sea mucho más grande en este marco que en su oponente y que en P&P el aprendizaje se represente con una maquinaria mucho más simple y económica; el espacio de gramáticas es mucho mayor en Optimalidad, aspecto que no es reflejado por sus defensores, que equiparan el espacio de gramáticas producto de combinar parámetros con el espacio resultado de combinar un conjunto de restricciones de diferentes maneras; sin embargo, en una lengua no hay una única jerarquía, y en las jerarquías de diferentes fenómenos se usan restricciones diferentes. Por ello, en ambas teorías se predice que la gramática correcta se encuentra dentro del espacio de gramáticas posibles (lo que significa que fuera de los límites de la GU no hay ningún espacio, aspecto muy deseado para el aprendizaje), pero el número y la simplicidad en cada caso es diferente⁸.

* Por esta misma razón, no se acaba de entender bien que Lightfoot (1999) defienda un modelo basado en señales más abstractas que los disparadores usualmente considerados para los parámetros, debido a que de este modo unas pocas señales evitan un espacio de hipótesis demasiado grande, producida por las posibilidades de combinación paramétrica. De nuevo, se está obviando el

Los principios no necesitan ordenación, sino conjunción en cuanto a su fijación paramétrica, algo bien diferente de postular que las restricciones necesitan ordenación.

Nótese, además, la aseveración de Cook & Newson (1996: 111-2) entre otros, según la que el niño podría mostrar una adopción previa o inicial de algunos valores de parámetros, que se corresponderían con aspectos no marcados entendibles como plasmación del principio de Avaricia (el mismo que determina el conservadurismo del proceso, evitando todo cambio en la fijación de parámetros excepto en el caso de que algún dato no pueda ser analizado con la fijación producida en ese momento; cfr. por ejemplo, Hyams 1986). De nuevo, con esto en mente, el modelo de Optimalidad se presenta mucho más abstracto y complejo y, por ello, menos verosímil.

5.4. La cuestión de la evidencia negativa

A mi juicio, el problema más severo de la visión explicitada consiste en sus asunciones al respecto de los tipos de evidencia: como he señalado previamente, la positiva, pero al tiempo, evidencia negativa implícita que, frente a P&P, es ofrecida por los análisis del propio niño. En el modelo rival de TO, la inexistencia de una forma en cierto contexto de manera sistemática puede ofrecer evidencia negativa implícita al niño (cfr. nota 7), como puede suceder en la determinación de si una lengua es *pro-drop* o no, de manera que este tipo de evidencia depende de si el aprendiz encuentra o no algo en el dato al que accede.

Sin embargo, Optimalidad va bastante más allá, asumiendo que el niño es incapaz de aplicar la degradación de restricciones si no es mediante la comparación de los datos positivos (o los análisis subyacentes adecuados) con los demás competidores del conjunto candidato, que de esta forma se constituyen en evidencia negativa implícita ofrecida por el propio niño. Con las palabras de Tesar & Smolensky (1998: 236) (cfr. también pp. 256-7),

the learner is not informed about the correct ranking by positive data in isolation; the role of the competing candidates must be addressed. This fact is not a liability, but an advantage: a comparative theory gives comparative structure to be exploited. Each piece of positive evidence, a grammatical structural description, brings with it a body of implicit negative evidence in the form of the competing descriptions. Given access to Gen and the underlying form (contained in the given structural description), the learner has access to these competitors.

Como se aprecia, el aspecto central de Optimalidad en relación a la teoría lingüística, es aplicado también de manera central en adquisición. Pero esta aplicación no se antoja muy convincente.

En primer lugar, el aprendiz debe trabajar con un grupo de candidatos que, como hemos visto, no están bien definidos en cuanto a los criterios y los límites para que conformen ese sistema, puesto que la función GENERADOR no es restricta, y no se

logro alcanzado a lo largo de los diferentes modelos, que ha producido un resultado estrictamente finito en adquisición.

le presta ninguna atención. Pero además, en igual medida sorprende que se afirme que esos análisis ofrecen la evidencia negativa implícita al niño, que debe ser consciente de no sólo el candidato ganador, sino de aquellos otros candidatos que son semejantes al ganador, aunque menos óptimos. Nótese que esta evidencia implícita no depende de datos abiertos, sino de análisis subyacentes del niño, bien a partir de datos recibidos, bien dirigidos a la producción. En caso de que el análisis dirigido a la producción no coincidiera con el análisis de una forma abierta, sería lógico plantear una comparación de ambos, pero no lo es tanto hacer lo mismo en relación a los otros candidatos conformados por GENERADOR diferentes de los dos apuntados. Nótese, además, que teóricamente no se trabaja con ninguna derivación, ya que el conjunto candidato es generado directamente; sin embargo, esto sólo es en teoría, puesto que en realidad para obtener ese conjunto deben actuar reglas, huellas, etc., aplicadas en GENERADOR, lo cual supone que el niño usa unos mecanismos negados por la teoría. Todo esto apunta, otra vez, a que el modelo es bastante más complejo y, además, asume como centrales algunos supuestos a los que sin embargo no les presta atención.

El papel central de la evidencia negativa implícita hace difícil especificar por qué ciertas secuencias no son enunciadas por el niño, como fenómenos de isla, puesto que tales fenómenos no pueden ser explicados mediante evidencia negativa implícita: hay muchas secuencias buenas que el niño no oye, simplemente porque no las ha encontrado en los datos (lo cual podría ser la situación de esos fenómenos), con lo que falta la comparación en cualquier conjunto candidato. Por ejemplo, piénsese en experimentos del estilo de Crain & Nakayama (1987) o Crain & Thornton (1998): se intenta forzar un error del niño mediante un contexto propicio para ello, pero el niño no lo comete, o también en Crain (1991), casos de adquisición para la que no hay ninguna experiencia disponible. Si en Optimalidad todo se explica mediante un ejemplo y mediante sus competidores directos, ¿por qué hay secuencias que el aprendiz no dice a pesar de que no hay *input* adecuado como para conformar un conjunto candidato en ese sentido? Como ya he señalado (cfr. *infra*, y nota 7), la evidencia negativa implícita de P&P no alude a análisis subyacentes, sino simplemente a la ausencia de un elemento en un contexto determinado y a la posibilidad de fijar un parámetro a partir de ella (cfr. Chomsky 1981: 9). De nuevo, la visión de TO sobre este tipo de evidencia es problemática, frente a P&P.

De ahí que la afirmación de Tesar & Smolensky (1998: 257) según la que “these (...) requirements contribute significantly to the overall goal of a learnability account with requirements realistic for a human child” debe ser fuertemente cuestionada.

6. CONCLUSIONES

Este trabajo ha expuesto y analizado los aspectos centrales de la visión que de la adquisición lingüística plantea TO, especialmente relacionada con el enfoque de Tesar & Smolensky (1998). Al igual que el modelo de P&P, TO reconoce un problema lógico, consistente en que los datos se quedan cortos como para alcanzar el estado final o estable, por lo que deben ser postulados mecanismos innatos que suplan esas carencias. El equivalente de la Gramática Universal de P&P, conformada por principios y especificaciones de las posiciones de los parámetros, consiste en TO en

un conjunto universal de restricciones, CONJUNTO, ordenadas de manera diferente según la lengua en cuestión. La labor del aprendiz consiste precisamente en descubrir la ordenación que garantice el llegar al estado estable.

Los defensores de este modelo de aprendizaje defienden, en varios sentidos, que la resolución del problema lógico se presenta como superior a la de su rival, P&P. Sin embargo, en el apartado 5 he cuestionado algunos aspectos a mi juicio problemáticos como para defender esa superioridad: (I) en primer lugar, la falta de aplicación global de la teoría lingüística y, en un nivel más particular, la aparente inviabilidad de tratamiento de un fenómeno como el de los infinitivos conjugados gallegos, lo que significa que no hay procedimiento de aprendizaje en este caso, (II) la indefinición absoluta de la función GENERADOR, que complica computacionalmente mucho las cosas y a la que no se le dedica atención, (III) el espacio de gramáticas posibles, sensiblemente mayor en TO que en P&P, ya que una gramática no está compuesta por una sola jerarquía, y (IV) la cuestión de la evidencia negativa implícita, asumida como ofrecida por el propio niño por medio de la comparación entre los miembros del conjunto candidato. En suma, no parece tan óptimo el modelo de adquisición de esta teoría como se pretende, siendo más sencillo el modelo P&P.

En último extremo, la resolución del problema lógico en TO parece casar peor que la postulada por P&P con el innatismo del lenguaje, equivalente a su adquisición entendida como crecimiento de un órgano mental: "Language learning is not really something that the child does; it is something that happens to the child" (Chomsky 1988: 134). Esta caracterización está mucho mejor recogida por el modelo paramétrico que por el basado en la ordenación de restricciones.

REFERENCIAS BIBLIOGRÁFICAS

- ARCHANGELI, D. (1997): "Optimality Theory: An Introduction to Linguistics in the 1990s", en D. Archangeli & T. Langendoen (eds.), 1-32.
- ARCHANGELI, D. & T. Langendoen (eds.) (1997): *Optimality Theory. An Overview*, Malden, MA & Oxford, Blackwell.
- BAKER, C. L. (1979): "Syntactic theory and the projection problem", *Linguistic Inquiry*, 10, 533-581.
- BARBOSA, P., D. FOX, P. HAGSTROM, M. MCGINNIS & D. PESETSKY (eds.) (1998): *Is The Best Good Enough? Optimality and Competition in Syntax*. Cambridge, MA, MIT Press.
- BERWICK, R. & P. NIYOGI (1996): "Learning from triggers", *Linguistic Inquiry*, 27, 605-622.
- CHOMSKY, N. (1980): *Rules and Representations*. New York, Columbia University Press. [Trad. esp., Méjico, F.C.E., 1983].
- (1981): *Lectures on Government and Binding*. Dordrecht, Foris.
- (1986): *Knowledge of Language: Its Nature, Origins and Use*. New York, Praeger. [Trad. esp., Madrid, Alianza, 1989].
- (1988): *Language and Problems of Knowledge*. Cambridge, MA, MIT Press [Trad. esp., Madrid, Visor, 1990].
- CLARK, R. (1992): "The selection of syntactic knowledge". *Language Acquisition*, 2, 83-149.

- CLARK, R. & I. ROBERTS (1993): "A computational model of language learnability and language change". *Linguistic Inquiry*, 24, 299-345.
- COOK, V. & M. NEWSON (1996): *Chomsky's Universal Grammar. An Introduction*, 2nd. edition, Oxford, Blackwell.
- COWIE, F. (1999): *What's Within? Nativism Reconsidered*, New York, Oxford University Press.
- CRAIN, S. (1991): "Language acquisition in the absence of experience", *Behavioral and Brain Sciences*, 14, 597-650.
- CRAIN, S. & M. NAKAYAMA (1987): "Structure dependence in grammar formation", *Language*, 63, 522-543.
- CRAIN, S. & R. THORNTON (1998): *Investigations in Universal Grammar. A Guide to Experiments on the Acquisition of Syntax and Semantics*, Cambridge, MA, MIT Press.
- DRESHER, B. E. (1999): "Charting the learning path: cues to parameter setting", *Linguistic Inquiry*, 30, 27-67.
- DRESHER, B. E. & J. KAYE (1990): "A computational learning model for metrical phonology", *Cognition*, 34, 137-195.
- FODOR, J. D. (1998): "Unambiguous triggers". *Linguistic Inquiry*, 29, 1-36.
- FORREST, S. (1993): "Genetic algorithms: principles of natural selection applied to computation", *Science*, 261, 872-878.
- FREIDIN, R. (1994): "Conceptual shifts in the science of grammar: 1951-1992", en C. Otero (ed.), *Noam Chomsky: Critical Assessments*, v. I, t. I, Routledge, London & New York, 653-690.
- GIBSON, E. & K. WEXLER (1994): "Triggers". *Linguistic Inquiry*, 25, 407-454.
- GILBERS, D. & H. DE HOOP (eds.) (1998): *Lingua*, 104. Special Issue: Optimality Theory.
- GOLD, E. M. (1967): "Language identification in the limit". *Information and Control*, 10, 447-474.
- HORNSTEIN, N. & D. LIGHTFOOT (1981): "Introduction", en N. Hornstein & D. Lightfoot (eds.), *Explanation in Linguistics*, London, Longman, 9-31.
- HYAMS, N. (1986): *Language Acquisition and the Theory of Parameters*, Dordrecht, Reidel.
- LASNIK, H. (1989): "On certain substitutes for negative data", en R. Matthews & W. Demopoulos (eds.), *Learnability and Linguistic Theory*, Dordrecht & Boston, Kluwer, 89-105.
- LIEBERMAN, P. (1998): *Eve Spoke. Human Language and Human Evolution*, New York, W.W. Norton.
- LIGHTFOOT, D. (1982): *The Language Lottery. Toward a Biology of Grammars*, Cambridge, MA, MIT Press.
- (1991): *How to Set Parameters. Arguments from Language Change*, Cambridge, MA, MIT Press.
- (1999): *The Development of Language. Acquisition, Change and Evolution*, Oxford, Blackwell.
- LONGA, V. M. (1994): "The Galician inflected infinitive and the theory of UG", *Catalan Working Papers in Linguistics*, 4, 23-44.
- (1999a): *La restricción de las gramáticas accesibles. La resolución del problema de la proyección en el dominio transformacional*, Santiago de Compostela, Servicio de Publicacións, Colección Lucus-Lingua, 5.
- (1999b): "Reseña de D. Archangeli & T. Langendoen (eds.), *Moenia*, 5.

- (en prensa): “Teoría da Optimalidade e ordenación das restricións universais: presentación e tratamento dos infinitivos conxugados”. en R. Álvarez & D. Vilavedra (eds.), *Homenaxe ó Profesor Xesús Alonso Montero*, Santiago de Compostela, Servicio de Publicacións.
- MARCUS, G. (1993): “Negative evidence in language acquisition”, *Cognition*, 46, 53-85.
- MORAVCSIK, E. (1988), “Agreement and markedness”, en M. Barlow & C. Ferguson (eds.), *Agreement in Natural Languages. Approaches, Theories, Descriptions*, Stanford, C.S. L.I., 89-106.
- MORGAN J. (1986): *From simple input to complex grammar*, Cambridge, MA, MIT Press.
- NEWSON, M. (1998): “On the Nature of Inputs and Outputs: A Case Study of Negation”, en P. Barbosa et al (eds.), 315-336.
- PETERS, S. (1972): “The projection problem: how is a grammar to be selected?”, en S. Peters (ed.), *Goals of Linguistic Theory*, New Jersey, Prentice-Hall, 171-188 [Trad. esp., Madrid, Gredos; 1983]
- PIATELLI-PALMARINI, M. (1989): “Evolution, selection and cognition: from ‘learning’ to parameter setting in biology and in the study of language”, *Cognition*, 31, 1-44.
- PINKER, S. (1997): *How the Mind Works*, New York, W.W. Norton.
- PRINCE, A. & P. SMOLENSKY (1993): *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms. Rutgers Optimality Archive (en prensa. Cambridge, MA, MIT Press).
- PULLEYBLANK, D. & W. TURKEL (1998): “The Logical Problem of Language Acquisition in Optimality Theory”, en Barbosa et al. (eds.), 399-420.
- RAPOSO, E. (1987): “Case Theory and Infl-to-Comp: The Inflected Infinitive in European Portuguese”, *Linguistic Inquiry*, 18, 85-109.
- RIGAU, G. (1995): “The properties of the temporal infinitive constructions in Catalan and Spanish”, *Probus*, 7, 279-301.
- SAMPSON, G. (1997): *Educating Eve. The ‘language instinct’ debate*, London, Cassell.
- SMOLENSKY, P. (1996): “On the comprehension/production dilemma in child language”, *Linguistic Inquiry*, 27, 720-731.
- TESAR, B. (1998a): “Error-driven learning in Optimality Theory via the efficient computation of optimal forms”, en P. Barbosa et al. (eds.), 421-435.
- (1998b), “An iterative strategy for language learning”, *Lingua*, 104, 131-145.
- TESAR, B. & P. SMOLENSKY (1998): “Learnability in Optimality Theory”, *Linguistic Inquiry*, 29, 229-268.
- TURKEL, W. (1996): “Acquisition by a GA-based model in spaces with local maxima”, *Linguistic Inquiry*, 27, 350-355.
- URIAGEREKA, J. (1998): *Rhyme and Reason. An Introduction to Minimalist Syntax*, Cambridge, MA, MIT Press.
- WEXLER, K. (1991): “On the argument from the poverty of stimulus”, en A. Kasher (ed.), *The Chomskyan Turn*, Oxford, Blackwell, 252-270.
- WEXLER, K. & P. CULICOVER (1980): *Formal Principles of Language Acquisition*, Cambridge, MA, MIT Press.
- WEXLER, K. & H. HAMBURGER (1973): “On the insufficiency of surface data for the learning of transformational languages”, en K. Hintikka, J. Moravcsik & P. Suppes (eds.), *Approaches to Natural Language*, Dordrecht, Reidel, 167-179.
- WILSON, E. O. (1999): *Consilience. La unidad del conocimiento*, Barcelona, Galaxia Gutenberg.