

¿Tiene la autoconciencia un fundamento lingüístico? Ernst Tugendhat y la Escuela de Heidelberg

Is Self-Awareness Linguistically Based? Ernst Tugendhat and the Heidelberg School

José Luis LÓPEZ DE LIZAGA
(Universidad de Zaragoza)

Recibido: 06/12/2012

Aceptado: 19/03/2013

Resumen

La aclaración de la estructura de la autoconciencia es crucial en el debate entre los partidarios del “giro lingüístico” y los defensores de la “filosofía de la conciencia”, pues si quedase probado que la autoconciencia tiene un fundamento lingüístico, el lenguaje se infiltraría hasta la raíz misma de la subjetividad. Teniendo presente esta pregunta fundamental, este artículo examina la polémica entre Ernst Tugendhat y los autores de la Escuela de Heidelberg (Dieter Henrich y Manfred Frank). Tugendhat concibe la autoconciencia como un fenómeno esencialmente lingüístico, mientras que la Escuela de Heidelberg detecta en la teoría de Tugendhat las dificultades de la concepción clásica de la autoconciencia como reflexión. Tras estudiar los argumentos de una y otra parte, el artículo concluye señalando la simetría de ambos enfoques, e indica brevemente las tareas que debería resolver una teoría alternativa: la distinción de diversos niveles de autoconciencia, y la aclaración de la relación entre el nivel prelingüístico y el nivel conceptual.

Palabras clave: autoconciencia, lenguaje, reflexión, Tugendhat, Henrich, Frank, Fichte.

Abstract

Clarifying the structure of self-awareness is a key point in the controversy between supporters of the “linguistic turn” and followers of the “philosophy of cons-

ciousness”, because as far as self-awareness is a linguistic phenomenon, language gets to the root of subjectivity. In the light of this fundamental issue, this paper examines the debate between Ernst Tugendhat and the Heidelberg School (Dieter Henrich and Manfred Frank) on self-awareness. Tugendhat understands self-awareness as an essentially linguistic phenomenon, while the Heidelberg School finds in Tugendhat’s theory the same problems that can be found in the classical reflexive theory of self-awareness. After considering both theories, this paper shows the unexpected symmetry between them and briefly points to two open issues an alternative approach should solve – namely, discriminating various levels in self-awareness and clarifying the relation between the pre-linguistic and the conceptual level.

Keywords: self-awareness, language, reflection, Tugendhat, Henrich, Frank, Fichte.

Durante algún tiempo parecía evidente que el “giro lingüístico” en filosofía se había consumado y era irreversible. Parecía probado que la conciencia no es el fundamento último de la constitución del mundo, ni de nuestro conocimiento de él, porque todas sus operaciones se realizan con medios lingüísticos.¹ Ésta era la premisa básica de un nuevo paradigma filosófico que fue abriéndose paso a lo largo del siglo XX en corrientes tan distantes en otros aspectos como son la hermenéutica alemana, la filosofía analítica anglosajona o el estructuralismo francés. Para los partidarios de este nuevo paradigma, ignorar el giro lingüístico sería tan estéril como regresar a una ontología prekantiana o precartesiana. Pero hoy esta conclusión parece apresurada, porque no existe aún, ni mucho menos, un acuerdo general en torno al alcance y los logros del giro lingüístico. De hecho, los continuadores del paradigma anterior –el paradigma de la “filosofía de la conciencia”²– han llevado a cabo un eficaz contraataque, y han intentado mostrar que la reflexión sobre el lenguaje no es capaz de desalojar a la conciencia de su antigua posición.

Esta disputa acerca del primado filosófico de la conciencia o del lenguaje está implícita en uno de los debates más interesantes de la filosofía alemana contemporánea, referido a la naturaleza de la autoconciencia. En un libro de 1979 titulado *Autoconciencia y autodeterminación*, Ernst Tugendhat intentaba mostrar que la autoconciencia tiene un fundamento lingüístico: somos conscientes de nosotros

¹ Adoptamos aquí la caracterización del giro lingüístico propuesta por Michael Dummett: consiste en la hipótesis de que “puede obtenerse una explicación [account] filosófica del pensamiento mediante una explicación filosófica del lenguaje”. Cf. Dummett, M.: *Origins of Analytical Philosophy*, Cambridge (Mass.), Harvard University Press, 1996, p. 4.

² Sobre esta periodización de la historia de la filosofía, cf. Apel, K.-O.: “¿Es posible actualmente un paradigma posmetafísico de filosofía primera?” en *Semiótica trascendental y filosofía primera*, Madrid, Síntesis, 2002.

mismos en la medida en que somos hablantes de un lenguaje, y más en concreto, en la medida en que disponemos de la habilidad de emplear correctamente el pronombre “yo” y los predicados referidos a estados mentales. Tugendhat formula esta teoría como una alternativa a la teoría clásica de la reflexión, pero también en confrontación con la teoría de la subjetividad desarrollada por la llamada “Escuela de Heidelberg”, compuesta por Dieter Henrich y sus seguidores, especialmente Manfred Frank. Henrich se apoya en Fichte para defender la existencia de una conciencia de sí prerreflexiva y, por supuesto, también prelingüística. Sin duda, la controversia entre Tugendhat y la Escuela de Heidelberg ha producido importantes contribuciones a la teoría de la autoconciencia, pero tiene también interés por una razón más general, puesto que el análisis de la autoconciencia puede considerarse casi como *experimentum crucis* en la elección entre el paradigma de la filosofía de la conciencia y el giro lingüístico. Si el lenguaje alcanzase incluso al *cogito* cartesiano, si pudiera probarse que incluso la más fundamental e indudable de nuestras certezas depende de nuestra competencia lingüística, entonces sería inevitable conceder al lenguaje el primado epistemológico sobre la conciencia. Si el *cogito* fuese un saber lingüístico, los partidarios del giro lingüístico habrían triunfado sobre los filósofos de la conciencia.

¿Logra la teoría de Tugendhat inclinar definitivamente la balanza a favor del giro lingüístico y contra la filosofía de la conciencia? En las páginas que siguen intentaré mostrar que la respuesta a esta pregunta es negativa. En primer lugar, expondré brevemente las dificultades de la teoría de la reflexión, de la que tanto Tugendhat como la Escuela de Heidelberg pretenden distanciarse (I). A continuación presentaré la teoría lingüística de la autoconciencia relacionándola con algunos aspectos de la filosofía del lenguaje de Tugendhat, en concreto su teoría de los términos singulares (II y III). Posteriormente examinaré las críticas de la Escuela de Heidelberg, cuyo objetivo es mostrar que el enfoque de Tugendhat reproduce las dificultades de la teoría de la reflexión: tal como sucede en esta teoría, las formas de autoconciencia basadas en el lenguaje remiten a un sustrato de autoconciencia más fundamental, que es tanto prerreflexivo como prelingüístico (IV). Estas críticas permiten concluir que Tugendhat proporciona el marco conceptual adecuado para la descripción de las formas más complejas de saber de sí, consistentes en la auto-atribución de propiedades mentales, pero que su enfoque es insuficiente para describir *todas* las formas de autoconciencia. No obstante, finalmente intentaré mostrar también que aceptar alguna forma de autoconciencia prelingüística no nos compromete con una misteriosa intuición inanalizable como la que los autores de la Escuela de Heidelberg hacen valer contra Tugendhat (V). El resultado de esta polémica sugiere, más bien, la necesidad de explorar algún otro enfoque teórico que sea capaz de articular distintos *niveles* de autoconciencia y de aclarar la relación entre los niveles prelingüísticos y los niveles dependientes del lenguaje (VI).

I

Uno de los modos en que tenemos conciencia de nosotros mismos es la *reflexión*, que Locke, por ejemplo, define como la “observación” introspectiva de “las operaciones internas de nuestra mente”.³ Puedo percibir un objeto, pero también puedo dirigir mi atención sobre mi percepción del objeto, y en este segundo caso llevo a cabo un acto de reflexión que me da noticia no ya (o no sólo) del objeto percibido, sino de mi propio acto de percibirlo. La autoconciencia así concebida es una conciencia intencional cuyo objeto es o bien el propio sujeto que lleva a cabo el acto, o bien otro episodio mental perteneciente al mismo sujeto.⁴ Nada parece más obvio e indiscutible que este modelo reflexivo en la descripción de la autoconciencia, y sin embargo en la filosofía contemporánea ha llegado a imponerse la tesis de que la reflexión no es el único modo de autoconciencia, ni tampoco –y esto es más importante– el modo más fundamental. Al contrario: la autoconciencia reflexiva sólo es inteligible sobre la base de algún otro género de saber de sí.

Dieter Henrich ha hallado en Fichte los argumentos definitivos contra la teoría de la autoconciencia como reflexión.⁵ De acuerdo con Fichte, esta teoría fracasa por un doble motivo: o bien sólo alcanza a ofrecer una explicación *circular* del fenómeno de la autoconciencia, es decir, una explicación que presupone aquello que debería explicar; o bien incurre en un *regreso infinito*, igualmente incapaz de dar cuenta de la autoconciencia. Comencemos por esta última dificultad. La concepción reflexiva supone que un sujeto inconsciente de sí cobra conciencia de sí mismo mediante un acto de reflexión. Pero es dudoso que la autoconciencia pueda explicarse mediante esta superposición de un acto reflexivo sobre un acto previo no reflexivo y no autoconsciente. Pues si el acto de reflexión por el que el yo cobra conciencia de sí no es autoconsciente, entonces requiere otro acto de reflexión para hacerse consciente de sí; un acto que a su vez necesitaría otro, y así sucesivamente. Fichte expone esta dificultad del siguiente modo:

Eres consciente de ti mismo como de aquello de lo que eres consciente [*Bewusstes*] sólo en la medida en que eres consciente de ti mismo como siendo consciente [*Bewusstseiendes*]; pero entonces lo que es consciente es a su vez aquello de lo que se es consciente, y debes hacerte consciente del ser consciente de eso de lo que eres consciente, y así *in infinitum*: y así no podrás alcanzar una primera conciencia.⁶

³ Locke, J.: *Ensayo sobre el entendimiento humano*, II.i.2, México, FCE, 1994, p. 83.

⁴ Tugendhat, E.: *Selbstbewusstsein und Selbstbestimmung*, Frankfurt, Suhrkamp, 1979, cap. 2; Zahavi, D.: *Self-Awareness and Alterity*, Evanston, Northwestern UP, 1999, p. 15.

⁵ Henrich, D.: *Fichtes ursprüngliche Einsicht*, Frankfurt, Klostermann, 1967; Henrich, D.: “Fichtes «Ich»”, en *Selbstverhältnisse*, Stuttgart, Reclam, 1982, pp. 62 y sigs. Henrich sostiene que la totalidad de la obra de Fichte puede reconstruirse como un intento (fracasado una y otra vez) de elaborar una descripción correcta de la autoconciencia tras haber rechazado el modelo reflexivo.

⁶ Fichte, J. G.: *Versuch einer neuen Darstellung der Wissenschaftslehre*, en M. Frank (ed.), *Selbstbewusstseinstheorien von Fichte bis Sartre*, Frankfurt, Suhrkamp, 1991, p. 19.

La teoría de la reflexión conduce, pues, a un regreso infinito. Pero para Fichte no es ésta la única dificultad a la que se enfrenta esta teoría: también está su carácter *circular*. En efecto, según la concepción reflexiva, el sujeto que reflexiona sobre sí mismo se descubre a sí mismo *como un sujeto*. Pero esto supone que la subjetividad, el yo o la autoconciencia ya están dados *antes* del acto de reflexión que, en principio, debería constituirlos, de modo que la autoconciencia no se origina en la reflexión, sino que la precede:

Se afirma que tu Yo sólo surge mediante el regreso [*Zurückgehen*] de tu pensamiento a sí mismo. En un pequeño rincón de tu alma aparece una objeción: *pienso*, pero antes de poder pensar debo *ser*; o bien esta otra objeción: me pienso, regreso a mí mismo; pero lo pensado, o aquello a lo que se regresa, debe ser antes de ser pensado o de que se regrese a ello.⁷

A primera vista, esta objeción puede parecer artificiosa y no muy convincente, pero su importancia queda más clara si se examinan algunas implicaciones paradójicas de la concepción reflexiva. Si el sujeto llega a saber de sí mismo por medio de la reflexión, entonces la autoconciencia es un *reconocimiento* de sí, un saber que enfrenta al sujeto con un objeto con el que se sabe *idéntico*. Pero o bien el sujeto que reflexiona ya se sabe idéntico con ese objeto que encuentra en el acto de reflexión, o bien el conocimiento de esa identidad de sujeto y objeto es el resultado del propio acto de reflexión. Pues bien, lo cierto es que cualquiera de estas dos alternativas arruina la teoría reflexiva de la autoconciencia. Si suponemos que el sujeto que lleva a cabo la reflexión ya se sabe idéntico con el objeto observado en ese acto, en realidad estamos suponiendo que el sujeto que lleva a cabo el acto de reflexión *ya es consciente de sí mismo*, de manera que la reflexión no funda la autoconciencia, sino que más bien la presupone. Pero supongamos ahora que, al contrario, el sujeto no sabe de sí mismo *antes* de realizar el acto de reflexión. En tal caso, resulta incomprensible cómo el acto de reflexión podría sustituir la diferencia entre sujeto y objeto por una relación de identidad; o por decirlo de otro modo: resulta incomprensible cómo sé que el episodio mental que convierto en mi objeto cuando reflexiono es un episodio mental *mío*, o cómo sé que esa subjetividad que aparece en el acto de reflexión es precisamente *la mía*. De manera que, o bien la reflexión presupone ya alguna forma (prerreflexiva) de autoconciencia, y por tanto su explicación de este fenómeno es una explicación circular, o bien la reflexión no es capaz en absoluto de explicar la autoconciencia.⁸

En su interpretación de estos argumentos fichteanos, el fenomenólogo Dan Zahavi señala que el problema fundamental de la teoría de la reflexión estriba en la

⁷ Id., p. 17.

⁸ Henrich, D.: "Fichtes «Ich»", op. cit., pp. 62-64.

concepción de la relación del sujeto consigo como una relación de intencionalidad: dado que la intencionalidad es la relación entre la conciencia y los objetos, es inevitable malinterpretar de entrada el fenómeno de la autoconciencia si se lo describe en estos términos.⁹ Las paradojas de la teoría de la reflexión serían simplemente las consecuencias de este error de enfoque que asimila conceptualmente la autoconciencia a la conciencia de objetos. Ahora bien, para corregir este error cabe optar entre, al menos, dos alternativas. Por un lado, podríamos renunciar al marco conceptual de la conciencia intencional y de la reflexión, e intentar describir la autoconciencia en términos de una conciencia no intencional y prerreflexiva. Ésta es la vía que escoge la Escuela de Heidelberg. En cambio, Tugendhat también rechaza el modelo de la reflexión, pero no rechaza el marco conceptual básico de la conciencia intencional, sino que más bien lo reinterpreta en términos lingüísticos: toda conciencia intencional es una conciencia *proposicional*, y la autoconciencia es la conciencia proposicional de que el sujeto se encuentra en un determinado estado psíquico. Ahora bien, es posible que, al mantener el esquema conceptual básico de la intencionalidad, en realidad esta revisión lingüística no logre sustraerse al problema de la circularidad de la teoría clásica de la autoconciencia como reflexión. Esto es lo que debemos examinar a continuación.

II

La teoría de la autoconciencia de Tugendhat se inscribe en un proyecto filosófico más amplio y ambicioso: la fundamentación del giro lingüístico, o la demostración del primado epistemológico del lenguaje sobre la conciencia. Este proyecto se desarrolla sobre todo en las importantes *Lecciones de introducción a la filosofía analítica del lenguaje* (1976), obra en la que Tugendhat defiende la tesis de que no existe una conciencia prelingüística, una referencia prelingüística a objetos. Para Tugendhat, la idea misma de una referencia prelingüística presupone que la operación básica de la conciencia es la *representación*, en el sentido de “tener (un objeto)-conscientemente-ante-sí”,¹⁰ no literalmente (como sucede en la percepción visual), sino en el sentido metafórico en que los filósofos de la conciencia hablan de la representación “no sensible” de estados de cosas, o también de la “percepción interna” del propio sujeto en la autoconciencia. Por más que este concepto de representación sea absolutamente central en el paradigma de la filosofía de la conciencia, en realidad se trata de un concepto insostenible. Tugendhat intenta probar que tener conciencia de un objeto no consiste en que algo se “haga presente” a la conciencia, sino más bien en *referirse* a un objeto mediante algún signo lingüístico: “no

⁹ Zahavi, D.: *Self-Awareness and Alterity*, op. cit., p. 16.

¹⁰ Tugendhat, E.: *Vorlesungen zur Einführung in die sprachanalytische Philosophie*, Frankfurt, Suhrkamp, 1976, p. 86 y sigs.

nos representamos los objetos, nos referimos a ellos”.¹¹ Así entendida, la conciencia de objetos es inseparable del lenguaje.

Para nuestros propósitos no necesitamos analizar en su totalidad la compleja argumentación desarrollada por Tugendhat en las *Lecciones* de 1976. Nos bastará con comprender su teoría de los *términos singulares*, o expresiones que nos permiten referirnos a entidades individuales. En relación con los términos singulares, Tugendhat sostiene una posición que podemos caracterizar como *holismo semántico*: sólo comprendemos el significado de una de estas expresiones en la medida en que sabemos emplear correctamente otras expresiones de función equivalente. Como veremos, Tugendhat extrae de este holismo semántico importantes consecuencias epistemológicas.

Tugendhat rechaza que los nombres propios sean el paradigma de los términos singulares. Pese a su aparente trivialidad, y a su éxito en la tradición de la filosofía de la conciencia,¹² la concepción de los términos singulares basada en el modelo de los nombres propios es incorrecta, pues la función de los términos singulares no consiste en *representar* directamente un objeto, o en estar en lugar de un objeto, sino más bien en *destacar* un objeto de una pluralidad: al emplear un término singular, “no sólo se dice «cuál», sino «cuál de todos»”¹³ los objetos es precisamente aquel al que nos referimos, esto es, aquel del que predicamos una determinada propiedad. Tugendhat llama *especificación* a esta función de los términos singulares, y extrae de ella consecuencias interesantes, como es la observación de que todo referirse a un objeto (en el sentido de especificarlo) supone la existencia de *otros* objetos: la especificación de un objeto no puede llevarse a cabo si no podemos identificar el objeto distinguiéndolo al mismo tiempo de todos los otros objetos a los que no nos referimos.¹⁴ Y es esencial comprender que esta identificación y distinción sólo es posible mediante una combinación de las distintas clases de términos singulares: expresiones deícticas, nombres propios y descripciones definidas (especialmente las descripciones localizadoras).¹⁵ Existe, pues, una remisión mutua entre los distintos términos singulares, y es esta remisión la que nos permite calificar como *holista* la teoría del significado de estas expresiones:

Si alguien quisiera saber cuál es la montaña más alta, y le llevásemos con los ojos vendados ante el Everest o a su cima, le quitásemos entonces la venda y le dijésemos “es esta de aquí”, sería muy natural que replicase: “¿pero cuál es la montaña que estoy viendo ahí?” Normalmente, una pregunta de este tipo se respondería con el nombre de la

¹¹ Id., 88.

¹² Id., p. 344 y sigs. Tugendhat toma a J. S. Mill como representante de este enfoque.

¹³ Id., p. 369. Tugendhat sigue a Strawson en esta concepción de los términos singulares.

¹⁴ Id., p. 371

¹⁵ Id., p. 347.

montaña. Pero el hecho de que la respuesta “la montaña más alta (o bien: la montaña que ves ahí delante) es el Everest” se admita normalmente como respuesta definitiva, sólo se debe a que se presupone que está claro *cuál* es la montaña así llamada, y con la pregunta de “cuál es” queremos decir dónde se encuentra, en qué relaciones espaciales se encuentra con otros objetos, con otros datos geográficos. Sólo ante una respuesta de este tipo parece absurdo seguir preguntando: “¿y cuál es la montaña que está en tal o cual lugar?” [...] Si ha de ser una auténtica identificación, la identificación demostrativa presupone por su parte la identificación espacio-temporal, no demostrativa.¹⁶

Este ejemplo muestra que, más que en los nombres propios o incluso en las expresiones deícticas, la identificación de un objeto mediante términos singulares se basa en su localización espacio-temporal. Referirse a un objeto perceptible significa, pues, identificarlo mediante su localización en un marco espacio-temporal en el que se inscriben también todos los otros objetos. Pero el ejemplo muestra también algo más importante, y es la circunstancia de que los términos singulares remiten constantemente unos a otros. Sólo podemos decir que hemos identificado un objeto cuando podemos referirnos a él como *el mismo* objeto mediante *las diferentes clases de términos singulares*: nombres propios, expresiones objetiva y subjetivamente localizadoras, deícticos. “No hay términos singulares de cualquier tipo que no remitan a términos singulares de otro tipo”.¹⁷ Esta remisión de los términos singulares es crucial, pues Tugendhat extrae de ella la consecuencia de que no podemos referirnos a objetos si no es mediante signos. No existe, pues, un contacto “directo” de los signos con los objetos; no existe algo tal como una referencia no lingüística a los objetos; no existe una capacidad prelingüística o a-lingüística de “mentar” o “representarse” objetos, como la que tradicionalmente ha admitido la filosofía de la conciencia. En realidad, esa capacidad inexistente es indisoluble de la concepción tradicional de la función de los términos singulares, que se orienta por los nombres propios. En efecto, si se afirma que estas expresiones están “en lugar de” los objetos, como si se tratase de etiquetas colocadas sobre ellos, hay que admitir también la posibilidad de referirse a los objetos sin la mediación de esos signos, de esas etiquetas, pues de otro modo no se explicaría cómo podríamos, precisamente, asociar el signo al objeto, al que tenemos que poder referirnos independientemente del signo para poder designarlo, nombrarlo o “etiquetarlo”: “la teoría del signo tradicional característica era una teoría del signo como representante: el signo representa [*vertritt*] algo que también podría darse sin el empleo de este o aquel signo –precisamente, en la representación”.¹⁸ Pero una concepción más adecuada de los términos singulares permite prescindir de esta idea de la representación

¹⁶ Id., p. 400.

¹⁷ Id., p. 473.

¹⁸ Id., p. 477.

prelingüística. En efecto, desde el momento en que interpretamos la función de los términos singulares como una función de *especificación*, entendiendo por tal la función de destacar un objeto de entre una pluralidad, queda rechazada la simple asociación de un signo a un objeto. Para referirnos a un objeto siempre tenemos que referirnos también de algún modo a la pluralidad de los otros objetos, puesto que no podemos llevar a cabo la especificación de un objeto sin distinguirlo de otros. Y esto sólo podemos hacerlo si dominamos el *sistema completo* de los términos singulares y somos capaces de referirnos a un mismo objeto mediante expresiones diferentes. “Toda pregunta «¿A qué objeto se refiere ‘a’?» encuentra su respuesta en *otro signo* que es sustituible por ‘a’”.¹⁹ Que seamos capaces de “representarnos” o “mentar” un objeto significa, pues, sencillamente esto: que podemos referirnos a ese objeto mediante términos singulares de distinto tipo.²⁰ Ahora estamos en condiciones de comprender la principal conclusión de esta investigación de Tugendhat: “no hay una referencia a un objeto –a un individuo– que esté exenta de signos”.²¹

III

El resultado de la investigación de Tugendhat sobre los términos singulares ha de ser aplicable también a la autoconciencia, y de su aplicación espera Tugendhat obtener una alternativa a la teoría de la autoconciencia como reflexión. En efecto, también Tugendhat rechaza el modelo reflexivo, pero no por los argumentos de Fichte –retomados luego por la Escuela de Heidelberg– que hemos expuesto más arriba, sino por considerar que la noción de “reflexión” se encuentra lastrada por la ficticia capacidad de “representación” que caracteriza en general a la filosofía de la conciencia. La noción de “reflexión” es una metáfora que toma su sentido de la percepción visual y lo extiende a un terreno enteramente inadecuado. Ni disponemos de algo tal como un “sentido interno”, ni la noción misma de una conciencia intencional, “dirigida a” un objeto, tiene un sentido preciso más allá de la percepción visual:

¹⁹ Id., p. 480.

²⁰ Id., p. 481.

²¹ Id., p. 482. Naturalmente, contra esta conclusión cabría objetar que la semántica de Tugendhat, por aguda y profunda que sea, no deja de ser una teoría sobre nuestro modo de emplear ciertos *signos*, pero nada nos impide seguir suponiendo una relación cognoscitiva (mención, representación) no lingüística con los objetos. Incluso si rechazamos toda enigmática capacidad prelingüística de “representación”, podemos seguir admitiendo la capacidad prelingüística de *percibir* objetos. Es más: sin esa capacidad no se comprende bien cómo podríamos “especificar” un objeto, destacarlo (como objeto percibido) de una pluralidad de objetos (asimismo percibidos). No obstante, dejaremos aquí de lado esta objeción.

Intenten por una vez mirar dentro de sí mismos. ¿No es también esto una metáfora inadecuada, como si pudiéramos reorientar también hacia dentro nuestra mirada, que está dirigida hacia fuera, en un así llamado acto de reflexión? (...) Dirán ustedes que se comprende en qué sentido la conciencia de algo es un estar dirigido. (...) Quizás nos acerquemos más a la cuestión si pensamos en la mirada que está dirigida a algo. Esto tiene sentido, y cuando miro algo, tengo conciencia de ello. Pero que hablemos aquí, a propósito de la percepción visual, de un “estar dirigido”, se debe simplemente al carácter espacial de esta percepción, al hecho de que está dirigida espacialmente desde un punto de vista. Si eliminamos lo espacial de este “estar dirigido”, tampoco queda nada del “estar dirigido”.²²

La alternativa que Tugendhat propone al modelo reflexivo de una conciencia intencionalmente dirigida hacia sí misma extiende a la descripción de la autoconciencia el modelo del saber lingüístico de los estados de cosas “externos”. Y esto implica que la autoconciencia es una forma de saber proposicional, y que su estructura reproduce –si bien con importantes diferencias– la estructura general “S es P” de toda proposición enunciativa. En efecto, si para comprender la conciencia analizamos las expresiones lingüísticas comunes, descubrimos que el objeto de toda conciencia intencional es una proposición o un estado de cosas:

Cuando consideramos relaciones de conciencia tales como desear, mentar, saber, tener la intención de hacer, temer, etc., constatamos que su objeto gramatical no es nunca una expresión que designa un objeto común, un objeto espacio-temporal, sino que su objeto gramatical es siempre una proposición nominalizada. No se puede desear, saber, etc. objetos espacio-temporales; cuando se desea, etc., siempre se desea que algo sea el caso.²³

Los objetos de la conciencia intencional no son entes espacio-temporales, sino estados de cosas o proposiciones. “Toda conciencia intencional en general es proposicional.”²⁴ O por decirlo de otro modo: la conciencia intencional no es conciencia de *x*, sino conciencia *de que (p)*.²⁵ Si aplicamos este modelo a la autoconciencia, ten-

²² Tugendhat, E.: *Selbstbewusstsein und Selbstbestimmung*, op. cit., p. 17.

²³ Id., p. 18.

²⁴ Id., p. 20. Incluso la mera representación de un objeto implica, según Tugendhat, una proposición, concretamente un juicio de existencia: aun si un objeto no existe “sólo es posible referirse conscientemente a un objeto *teniéndolo por* existente, y que un objeto existe es, por supuesto, una proposición.”

²⁵ Un tanto sorprendentemente, Tugendhat reconoce que la tesis de que toda conciencia intencional es conciencia de una proposición o estado de cosas no implica por sí misma que toda conciencia intencional sea una conciencia lingüística, ni que su objeto sea un objeto de naturaleza lingüística. Tugendhat admite que el estatus ontológico de los estados de cosas “es controvertido”, renuncia por su parte a intentar dilucidarlo, y admite que la posibilidad de identificar estados de cosas con independencia del lenguaje es una cuestión abierta, al menos en el caso de los animales y los niños pequeños. Cf. id., pp. 19-21.

dremos que admitir que también aquí se trata de un saber proposicional, un saber de estados de cosas: concretamente, el saber *de que* me encuentro en un determinado estado mental, o en un “estado φ ”, por emplear la terminología que escoge el propio Tugendhat. Y al igual que sucede con el saber acerca de estados de cosas intramundanos, también el saber en el que consiste la autoconciencia puede expresarse en enunciados del tipo “(yo sé que) S es P”, con la importante salvedad de que P corresponde a predicados φ , y S designa el sujeto que en cada caso emite el enunciado.

De acuerdo con Tugendhat, las semejanzas estructurales entre el saber de estados de cosas intramundanos y el saber de sí, o entre las proposiciones enunciativas en tercera y en primera persona, pueden reconocerse tanto en (1) el holismo semántico que afecta al sujeto gramatical de las proposiciones φ , como en (2) la función de *caracterización* que, al igual que cualquier predicado en cualquier persona gramatical, cumplen los predicados φ en primera persona. Veámoslo con más detalle.

1. La estrategia de Tugendhat para desembarazarse del modelo tradicional de la autoconciencia como “percepción interior” del “Yo” o de sus estados mentales consiste en atenerse metodológicamente a un análisis del significado de la palabra “yo”. Desde esta perspectiva metodológica, lo primero que podemos afirmar acerca de la expresión “yo” es que se trata de un pronombre personal, y que como tal pertenece, junto con los demostrativos o algunos adverbios de tiempo y lugar (como “aquí” y “ahora”) a la categoría de las expresiones deícticas, cuyo significado depende del contexto en que se emplean. El pronombre “yo” es la expresión con la que el hablante se designa a sí mismo como el sujeto de una proposición en la que se atribuye un predicado. Y tal como sucede con todas las expresiones deícticas –y con todos los términos singulares–, el pronombre “yo” forma un grupo con *otras* expresiones, de tal modo que el significado (es decir, las reglas de uso) de cada una de ellas depende sistemáticamente del significado de las demás. Por ejemplo, el adverbio “aquí” puede ser sustituido por “allí” si se designa el mismo lugar desde otro sitio, mientras que desde este lugar que llamamos “aquí”, todos los otros lugares se designan como “allí”. Pues bien, esta referencia sistemática de unas a otras expresiones se reproduce en el caso de los pronombres personales: el empleo de la expresión “yo” presupone en el hablante el conocimiento de las reglas de uso de las expresiones “tú”, “él”, etc. El empleo correcto del pronombre “yo” implica, pues, la existencia de una *pluralidad* de hablantes, así como la comprensión de que la persona que se refiere a sí misma mediante el pronombre “yo” es *la misma* que otros designan con los pronombres “tú” o “él”, es decir, la comprensión de que los hablantes “pueden referirse mutuamente los unos a los otros”.²⁶ El *holismo semántico* de los términos singulares se aplica también a los pronombres personales.

2. La semejanza estructural entre las proposiciones referidas a estados de cosas intramundanos y las proposiciones φ se reconoce también en los predicados.

²⁶ Id., p. 74-75.

Tugendhat propone atribuir a estas expresiones la función de *caracterizar* al sujeto de la proposición: una expresión predicativa como “rojo” no designa o representa algo (una propiedad universal), sino que su función consiste en *clasificar* un objeto y *distinguirlo* de otros.²⁷ Y lo que es más importante: al igual que sucede con los predicados referidos a propiedades físicas, los predicados ϕ caracterizan su objeto manteniendo exactamente el mismo sentido en primera persona y en tercera persona. Tugendhat se apoya en Strawson para afirmar un principio de *simetría veritativa* entre las proposiciones expresadas en primera persona y las proposiciones en tercera persona. Este principio es trivial cuando nos referimos a propiedades físicas, pero se vuelve más interesante si se aplica también, como hacen Strawson y Tugendhat, a los predicados y proposiciones ϕ . En su obra *Individuos*, Strawson sostiene que sólo es posible adscribirse estados de conciencia a sí mismo en la medida en que los adscribimos *exactamente en el mismo sentido* a otras personas.²⁸ Y Tugendhat adopta el mismo principio relacionándolo con la pluralidad de hablantes que es condición de posibilidad del empleo del pronombre “yo”:

Hemos visto que la aclaración y con ello el significado de la palabra ‘yo’ implica que quien dice ‘yo’ sabe que otras personas pueden referirse con ‘él’ o con ‘este hombre’ a ese al que él se refiere con ‘yo’. [...] Pues bien, esto tiene consecuencias para las proposiciones ϕ . Pues cuando un predicado conviene a una entidad, simplemente le conviene, y da igual con qué término singular es designado. Ésta es la llamada Ley de Leibniz, que se funda en el significado de ‘=’: si $a=b$, entonces todo predicado que convenga a a , conviene también a b . Y para las proposiciones ϕ , de aquí se sigue lo siguiente: *la proposición “yo ϕ ”, pronunciada por mí, es verdadera exactamente cuando es*

²⁷ Tugendhat, E: *Vorlesungen zur Einführung in die sprachanalytische Philosophie*, op. cit., pp. 182-183: “Cuando aplicamos un predicado a algunos objetos y no a otros, con ello *clasificamos* a los que lo aplicamos, y con ello los *distinguimos* al mismo tiempo de aquellos a los que no lo aplicamos. Cuando aplicamos un predicado a un objeto, declaramos que es tal como los otros objetos a los que aplicamos el predicado, y que no es como aquellos a los que no lo aplicamos, y esto significa que lo caracterizamos como tal. La función de caracterización consiste en clasificar-y-distinguir.” Esta función de caracterización puede prescindir del supuesto de entidades universales denotadas por las expresiones predicativas, y atenerse a una teoría del significado como uso, como la propuesta por Wittgenstein. Cf. id., p. 188: “El significado del predicado no consiste en la característica general; antes bien, ésta no aparece en absoluto en la explicación del significado”. Ahora bien, habría que demostrar que esta función de caracterización se distingue realmente –y no sólo terminológicamente– de la función de *nombrar* o representar propiedades universales. Tugendhat intenta hacerlo (en las Lecciones 11 y 12) mediante un largo diálogo imaginario entre un “conceptualista” y un “nominalista”; un diálogo en el que, sin embargo, no necesitamos detenernos aquí.

²⁸ Strawson, P. F.: *Individuos*, Madrid, Taurus, 1989, p. 101-102: “Una condición necesaria para que uno se adscriba estados de conciencia, experiencias, a sí mismo, de la manera en que lo hace, es que los adscriba también, o esté dispuesto a adscribirlos, a otros que no son él mismo. Esto significa [...] que las expresiones adscriptivas se usan exactamente en el mismo sentido cuando el sujeto es otro que cuando el sujeto es uno mismo.”

verdadera la proposición “ ϕ ”, cuando es pronunciada por alguien distinto de mí y que se refiere a mí cuando dice “él”.²⁹

De acuerdo con este principio, significan exactamente *lo mismo* el predicado de una proposición ϕ en primera y en tercera persona (por ejemplo: “estoy angustiado” y “él está angustiado”). A favor de esta identidad de significado habla también el análisis del modo en que aprendemos el significado de los predicados mentales, que Tugendhat toma de Wittgenstein. Tal como muestran las *Investigaciones filosóficas*, especialmente en los pasajes en los que Wittgenstein critica la posibilidad de un “lenguaje privado”, el aprendizaje de los términos referidos a lo mental, incluido el aprendizaje de su uso en *primera* persona, presupone la comprensión de su significado en *tercera* persona. Wittgenstein insiste en que los niños no aprenden el significado de los términos mentales “por introspección”, sino sustituyendo la expresión natural de ciertas vivencias por su expresión verbal convencional o *simbólica*. Cuando un niño siente dolor y llora, podemos interpretar que el llanto es la expresión natural del dolor. Posteriormente los niños aprenden a sustituir esa expresión natural por una expresión verbal, como es la proposición “me duele”, de tal manera que “las palabras se conectan con la expresión primitiva, natural, de la sensación y se ponen en su lugar”.³⁰ Esta expresión verbal no *describe* una sensación “privada”, sino que simplemente sustituye a una *conducta* no verbal: el llanto. Y así como no decimos que el llanto describe una sensación privada, tampoco podemos decir que lo haga la proposición “me duele”. De este modo rechaza Wittgenstein la hipótesis de que los términos mentales obtengan su significado mediante actos de ostensión referidos a los episodios mentales del hablante. Pero si rechazamos la introspección o la “percepción interna” como fuente de nuestro conocimiento del significado de los predicados ϕ , parece que ya sólo podemos recurrir a la observación, a la percepción “externa” de la conducta propia y ajena, para explicar a otros, y aprender nosotros mismos, el significado de esos términos. Y de hecho, Wittgenstein afirma que si lo mental no se inscribiera en un juego de lenguaje compartido intersubjetivamente, que incluye determinadas *conductas* públicamente observables y sus expresiones correspondientes, no sería posible en absoluto aprender el significado de las expresiones referidas a lo mental: “¿Cómo sería si los hombres no manifestasen su dolor (no gimiesen, no contrajesen el rostro, etc.)? Entonces no se le podría enseñar a un niño el uso de la expresión «dolor de muelas»”.³¹

Ahora bien, ¿significa todo esto que no existe *ninguna* diferencia entre el conocimiento de un estado de cosas intramundano y la conciencia de encontrarse en un determinado estado mental, como la que se expresa en una proposición del tipo “yo

²⁹ Tugendhat, E.: *Selbstbewusstsein und Selbstbestimmung*, op. cit., p. 88.

³⁰ Wittgenstein, L.: *Investigaciones filosóficas*, Barcelona, Crítica, 1988, §244, p. 219.

³¹ Id., §257, pp. 225-227.

sé que yo φ ”? Ciertamente, no. Las diferencias entre la expresión lingüística de nuestro saber de estados de cosas y la de nuestro saber de nuestros propios estados mentales son tan importantes como sus semejanzas estructurales. Y a decir verdad, la plausibilidad de la teoría de Tugendhat depende de si son más relevantes las semejanzas o las diferencias.

Tugendhat no pasa por alto dos diferencias muy destacadas entre el saber proposicional de estados de cosas intramundanos y el saber proposicional de sí, relacionadas con las funciones tanto de los términos singulares como de los predicados, esto es, con las funciones de *especificación* y *caracterización*. A diferencia de lo que sucede con cualesquiera otros términos singulares referidos a objetos (o con cualesquiera otros pronombres referidos a otras personas), el pronombre “yo” empleado en proposiciones φ es *inmune a todo error de identificación*. Esto se debe a que, al expresar una proposición en primera persona, no llevamos a cabo un acto de especificación del sujeto gramatical como el que se produce mediante el empleo de cualquier otro término singular.³² En una proposición como “siento dolor” no destacamos un objeto entre una pluralidad de objetos, al que después atribuimos una propiedad (sentir dolor), y al que además identificamos con nosotros mismos.³³ Este proceso de especificación simplemente no tiene lugar en las proposiciones φ en primera persona. Y por lo que respecta a la función de los predicados, es evidente que en las proposiciones en primera persona que contienen un predicado φ tampoco *caracterizamos* al sujeto gramatical del mismo modo en que lo hacemos en tercera persona. Consideremos las proposiciones “sé que estoy angustiado” y “sé que él está angustiado”: en un caso tenemos un saber *inmediato* y *cierto* de un estado φ , mientras que en el otro caso nuestro saber es sólo mediato y falible, puesto que a partir de ciertos rasgos conductuales observables inferimos –y siempre podemos equivocarnos– que el sujeto observado se encuentra en un estado φ .

El propio Tugendhat admite que el pronombre personal de primera persona no cumple una función de especificación, y completa su principio de simetría veritativa con un principio de *asimetría epistémica* entre las perspectivas de primera y la tercera persona: “los estados φ son sabidos [*gewusst*] (...) de modo diferente por el afectado y por los demás.”³⁴ De manera que, incluso aceptando la hipótesis de que la autoconciencia es un tipo de conciencia proposicional, parece inevitable admitir que ciertos rasgos de la conciencia de sí la distinguen de la conciencia de objetos o estados de cosas. Tugendhat analiza esas diferencias sin renunciar a su hipótesis de partida. Sin embargo, no está claro que, tras reconocer estas diferencias, pueda seguir sosteniendo el carácter esencialmente lingüístico de la autoconciencia.

³² Por esta razón, algunos autores niegan que la función de esta expresión consista realmente en *referirse* a entidad alguna. Cf. Zahavi, D.: *Self-Awareness and Alterity*, op. cit., p. 10 y sigs.; Lailach-Henrich, A.: *Ich und die Anderen*, Berlin, Gruyter, 2011.

³³ Zahavi, D.: *Self-Awareness and Alterity*, op. cit., p. 5.

³⁴ Tugendhat, E.: *Selbstbewusstsein und Selbstbestimmung*, op. cit., p. 89.

IV

De este punto arrancan las críticas de la “Escuela de Heidelberg”, cuyos representantes detectan en la teoría de la autoconciencia de Tugendhat la *circularidad* en que incurren todos los intentos de explicar la autoconciencia a partir de algún otro saber, en este caso la competencia lingüística en el empleo de los pronombres personales.³⁵ La semántica de Tugendhat, y en particular su interesante concepción holista de los pronombres personales, nos proporciona una teoría sobre nuestro lenguaje, pero no una teoría de la autoconciencia. No basta, en efecto, con emplear el pronombre “yo” para tener autoconciencia, y esta es la razón por la que no atribuimos autoconciencia a una máquina, un autómatas o una computadora diseñada para pronunciar este pronombre, incluso en una aparente conversación con un ser humano.³⁶ Para poder ser cualificado como un sujeto autoconsciente, el hablante que emplea el sistema de los pronombres personales debe disponer de algún otro saber, además de su capacidad de emitir la expresión “yo”. Sólo sobre la base de ese otro saber es posible emplear *con sentido* (y no de un modo simplemente mecánico, como lo haría una máquina o un loro) el sistema de los pronombres. La propia competencia lingüística en el empleo del pronombre “yo” presupone, pues, algún tipo de saber *diferente y previo* al saber lingüístico. Y ese otro saber sólo puede ser ya, precisamente, alguna forma de *autoconciencia*, pues sólo ésta nos permite identificar a las personas como tales, y diferenciarlas, por ejemplo, de las máquinas capaces de “hablar”:

Si examinamos el modo en que son identificadas las personas (...), tal identificación tendrá que referirse siempre también a aquello que las personas son en la medida en que puede reconocérseles una perspectiva sobre su propia vida y su comportamiento. Si las personas son captadas como hablantes, esta perspectiva interna será aquello que las haga capaces de emplear significativamente el lenguaje.³⁷

³⁵ Henrich, D.: “Noch einmal in Zirkeln: eine Kritik von Ernst Tugendhats semantischer Erklärung von Selbstbewusstsein” en C. Bellut *et al.* (eds.), *Mensch und Moderne*, Würzburg, Königshausen & Neumann, 1989, pp. 93-132. En una réplica de 2005 a las objeciones publicadas por Henrich en 1989, Tugendhat niega que su propósito en *Autoconciencia y autodeterminación* consistiese en proponer una *explicación* de la autoconciencia. Esto es cierto, si Tugendhat se refiere a una explicación genética como la que propone la psicología evolutiva. Pero no creo que esta réplica logre desmontar por sí sola las objeciones de la Escuela de Heidelberg, que se dirigen contra la tesis de que la autoconciencia es un fenómeno vinculado al lenguaje. Y esta tesis, sin duda, fue sostenida por Tugendhat en *Autoconciencia y autodeterminación*. Cf. Tugendhat, E.: “Über Selbstbewusstsein: Einige Missverständnisse”, en Th. Grundmann *et al.* (eds.), *Anatomie der Subjektivität*, Frankfurt, Suhrkamp, 2005, pp. 247-254.

³⁶ Henrich (“Noch einmal in Zirkeln”, *op. cit.*, p. 102) emplea este otro ejemplo: sólo decimos figuradamente que una locomotora nos avisa de su aproximación cuando emite un pitido de advertencia. Otro tanto sucede cuando atribuimos subjetividad a una máquina capaz de pronunciar la palabra “yo”.

³⁷ Henrich, D.: “Noch einmal in Zirkeln”, *op. cit.*, p. 103.

La autoconciencia es el saber del que depende el empleo correcto de los pronombres personales. “Sin ese saber, la emisión de «yo» sería un mero sonido, no un empleo significativo del lenguaje”.³⁸ Ahora bien, esto nos obliga a invertir por completo la teoría de Tugendhat: si la condición de posibilidad de un uso correcto del pronombre “yo” es, precisamente, la autoconciencia, entonces ésta no tiene un origen lingüístico ni depende esencialmente del lenguaje, sino que está más bien *supuesta* en nuestra competencia para emplear correctamente el sistema de los pronombres personales.³⁹ Así pues, aunque Tugendhat propone su enfoque lingüístico como alternativa a la concepción reflexiva de la autoconciencia, su propia teoría reproduce la circularidad de ésta: presupone el fenómeno que se propone explicar –la autoconciencia– entre las condiciones que deberían explicarlo –el empleo correcto del pronombre “yo”. Y de acuerdo con Henrich, si Tugendhat pretendiese evitar la objeción de circularidad reduciendo la autoconciencia al empleo del pronombre “yo”, sólo lograría que “se esfumase el fenómeno” que la teoría se proponía explicar.⁴⁰

Pero las dificultades de la teoría de Tugendhat no sólo aparecen en relación con las condiciones de un empleo significativo del pronombre “yo”, sino también en su teoría de las expresiones predicativas, concretamente en relación con el principio de *simetría veritativa* entre las proposiciones ϕ en primera y tercera persona. Pese a las observaciones de Wittgenstein sobre el aprendizaje del significado de los términos mentales, cabe argumentar que la relación entre ambas personas gramaticales es exactamente la inversa de la que afirma Tugendhat. No es posible que sólo aprendamos en tercera persona el significado de los predicados ϕ que luego nos atribuimos a nosotros mismos en primera persona, puesto que toda identificación de nosotros mismos en tercera persona presupone ya la primera persona como su condición de posibilidad. La posición de Henrich o Frank coincide en este punto con las aportaciones de un filósofo analítico como Sydney Shoemaker, quien ha agumentado a favor del primado de la perspectiva de primera persona sobre la tercera. Atribuirme a mí mismo una propiedad física, corporal –por ejemplo, el tipo de conducta que, según Wittgenstein, el niño observa en otros y asocia a determinados estados ϕ – significa afirmar que *mi* cuerpo tiene esa propiedad, pero al intentar aclarar lo que quiero decir cuando hablo de *mi* cuerpo –a diferencia de ese cuerpo de allí, o del cuerpo de ese otro hombre que observo, etc.–, tengo que recurrir inevitablemente a un lenguaje que ya incluye proposiciones ϕ en primera persona, tales como “veo”, “siento”, etc.:

³⁸ Id. p. 107.

³⁹ Y si la autoconciencia no sólo es anterior e independiente del lenguaje, sino de toda relación intersubjetiva, esta objeción contra el planteamiento de Tugendhat puede hacerse extensiva a otras teorías afines, como las de Mead o Habermas. Cf. sobre esto Frank, M.: *Selbstbewusstsein und Selbsterkenntnis*, Stuttgart, Reclam, 1991, p. 410 y sigs.

⁴⁰ Henrich, D.: “Noch einmal in Zirkeln”, op. cit., p. 128, p. 101.

Podríamos imaginarnos un pueblo que habla una lengua primitiva que contiene un pronombre de primera persona, pero no predicados P* [equivalentes de los predicados ϕ de Tugendhat en la terminología escogida por Shoemaker] (...) Sin embargo, cuando se intenta describir las circunstancias en las que pueden tener lugar las atribuciones a sí mismo y aportar las razones para tales atribuciones a sí mismo, hay que emplear necesariamente ciertos predicados, los predicados P*, que no pueden expresarse en el lenguaje que se nos ha presentado.⁴¹

Este argumento de Shoemaker parece probar el primado del lenguaje de lo mental sobre el lenguaje fisicalista, pero también es interesante para la teoría de la autoconciencia, puesto que implica que toda atribución a uno mismo de un predicado corporal en el mismo sentido en que lo atribuimos a una tercera persona presupone ya una conciencia de sí mismo en primera persona, no mediada por la perspectiva de tercera persona: “el conocimiento perceptivo de sí mismo presupone un conocimiento de sí no perceptivo, de modo que no todo conocimiento de sí puede ser perceptivo.”⁴² Al menos en un nivel muy básico, mi conciencia de mí mismo no consiste en (o no se reduce a) la caracterización de mis propios estados mentales mediante predicados cuyo significado he aprendido primero en tercera persona.

Y en realidad esta otra forma “no perceptiva” de conciencia de sí aparece, pese a todo, en *el interior* de la teoría de Tugendhat, en concreto en la distinción entre la *simetría veritativa* y la *asimetría epistémica* entre las perspectivas de primera y tercera persona. Para Manfred Frank, esta distinción en el fondo arruina la propuesta de Tugendhat. Intentemos comprender esta objeción de Frank mediante las siguientes consideraciones. De acuerdo con el principio strawsiano de *simetría veritativa*, el significado de un término ϕ en primera persona es exactamente el mismo que en tercera persona, mientras que de acuerdo con el principio de *asimetría epistémica* que Tugendhat añade al principio anterior, la atribución de un predicado ϕ presenta, como ya hemos indicado, varias características enteramente diferentes en primera y en tercera persona. Pero si desde la perspectiva de primera persona no identificamos al sujeto (yo mismo) entre una pluralidad de otros sujetos, ni podemos cometer errores de identificación, ni inferimos faliblemente la existencia del estado ϕ a partir de la observación de nuestro comportamiento; y en cambio en tercera persona necesitamos identificar al sujeto (como a cualquier objeto del mundo al que atribuimos una propiedad), y podemos errar en esa identificación, y además nuestra atribución del predicado ϕ es el resultado de una *inferencia* (meramente hipotética

⁴¹ Shoemaker, S.: “Selbstbezug und Selbstbewusstsein”, en M. Frank (ed.), *Analytische Theorien des Selbstbewusstseins*, Frankfurt, Suhrkamp, 1994, p. 58. Cf. también Frank, M.: *Selbstbewusstsein und Selbsterkenntnis*, op. cit., p. 39.

⁴² Shoemaker, S.: “Personal Identity”, en S. Shoemaker / R. Swinburne, *Personal Identity*, Oxford, Blackwell, 1984, p. 105.

y falible) a partir de la conducta corporal observada, ¿hasta qué punto es realmente posible afirmar una simetría veritativa entre ambas perspectivas?

La respuesta de Tugendhat a esta pregunta quizás podemos hallarla en una frase en la que comenta a Wittgenstein: “si nos imaginamos que tenemos un dolor, nos representamos cómo nos comportaríamos”.⁴³ Tugendhat insistiría, pues, en que las experiencias ϕ en primera persona toman su sentido de la observación de esas experiencias en tercera persona. Pero esta posible réplica de Tugendhat a la crítica de Manfred Frank al principio de simetría veritativa no parece suficiente para desmontar la objeción. Es verdad que asociamos los estados ϕ a determinadas conductas observables, pero sigue habiendo una diferencia entre *sentir* un dolor y *conocer el significado de la conducta* (verbal o no) que lo expresa, y esta diferencia no puede borrarse simplemente decretando la identidad del significado de los términos ϕ en primera y en tercera persona. Parece, pues, que la aplicación del principio de simetría veritativa a las proposiciones ϕ conduce a conclusiones falsas, si tenemos en cuenta las diferencias en cuanto a los funciones de especificación y caracterización de un sujeto en las proposiciones en primera y en tercera persona. Por eso Frank sostiene que, desde el momento en que el propio Tugendhat reconoce la asimetría epistémica entre una y otra perspectiva, ya no puede seguir sosteniendo coherentemente su principio de simetría veritativa, que queda reducido a un mero postulado no demostrado, a una petición de principio:

Desde un punto de vista epistemológico no está nada claro que lo que tengo en la autoconciencia de un modo no identificador y sin observación, sea lo mismo que aquello que, en la actitud adecuada, tengo ante los ojos de un modo sensorial. (...) Tugendhat se limita a suponer esta identidad, y lo hace con la vista puesta exclusivamente en la semántica. Este supuesto (...) tiene sentido desde el punto de vista de la semántica formal, pero no implica ninguna garantía de realidad. (...) Por tanto, la asimetría epistémica resulta ruinosa para la bella armonía de la construcción de Tugendhat, basada en un mero postulado de simetría semántica.⁴⁴

En conclusión, de acuerdo con la Escuela de Heidelberg la teoría lingüística de la autoconciencia propuesta por Tugendhat no es otra cosa que una variante de la teoría de la autoconciencia como reflexión. Es cierto que Tugendhat rechaza la hipótesis de una “percepción interna” por considerarla una mera metáfora incomprensible, pero inadvertidamente se atiene a otro supuesto básico del modelo reflexivo: la concepción de *toda* conciencia como conciencia intencional referida a un objeto. La única diferencia a este respecto entre la teoría de Tugendhat y la concepción tradicional de la autoconciencia como reflexión estriba en que para Tugendhat toda conciencia intencional es proposicional o está mediada por el lenguaje. Pero

⁴³ Tugendhat, E.: *Selbstbewusstsein und Selbstbestimmung*, op. cit., p. 125.

⁴⁴ Frank, M.: *Selbstbewusstsein und Selbsterkenntnis*, op. cit., pp. 427-428.

esta teoría lingüística reproduce la circularidad de la teoría reflexiva y, en nombre de un postulado semántico de identidad de significado, tiende a difuminar la diferencia epistemológica entre las perspectivas de primera y tercera persona. Pues bien, bastaría con abandonar el “totalitarismo de la doctrina de la proposicionalidad”,⁴⁵ pero también la relación intencional de sujeto-objeto, para allanar el camino de la comprensión de la autoconciencia. Y así, Henrich y Frank hacen valer contra Tugendhat el dato irrefutable –aunque también, como veremos, aparentemente inanalizable– de la conciencia de sí como una conciencia *no proposicional* y *no intencional*.

V

En mi opinión, las críticas de la Escuela de Heidelberg a Tugendhat permiten concluir que la teoría que hace de la autoconciencia un fenómeno vinculado al lenguaje no es una teoría incorrecta, pero sí *incompleta*, de la autoconciencia. Probablemente existen formas de autoconciencia que son inconcebibles sin el lenguaje, puesto que existen pensamientos inconcebibles sin lenguaje, y algunos de esos pensamientos se refieren al propio sujeto que los piensa. Wittgenstein muestra esto con el ejemplo de Ballard, un sordomudo citado por William James que afirmaba recordar los pensamientos sobre Dios y el origen del mundo que tenía *antes* de aprender “los rudimentos del lenguaje escrito.”⁴⁶ Con razón pone Wittgenstein en duda la veracidad de tales recuerdos, puesto que parece improbable tener conciencia de objetos tan complejos como Dios o el origen del mundo sin disponer del marco categorial que adquirimos junto con la competencia lingüística.⁴⁷ Sin embargo, por otra parte parece cierto, como pretende la Escuela de Heidelberg, que el empleo correcto del pronombre “yo” presupone alguna forma de autoconciencia más básica que aquella autoconciencia lingüística que consiste en atribuirse a sí mismo un predicado ϕ . Ahora bien, ¿qué más podemos decir acerca de esta otra forma de autoconciencia?

A decir verdad, los mayores logros de la Escuela de Heidelberg se localizan en la crítica del modelo reflexivo de la autoconciencia (y sus inesperadas variantes, como la propia teoría de Tugendhat). Henrich y Frank intentan evitar las dificultades de la teoría reflexiva por una vía muy distinta al enfoque lingüístico de

⁴⁵ Id., p. 226.

⁴⁶ Cf. Wittgenstein, L.: *Investigaciones filosóficas*, op. cit., §342.

⁴⁷ En otros casos (por ejemplo, un dolor no claramente localizado, o un estado de ánimo impreciso), podemos tener la experiencia, pero no podemos identificarla y reconocerla como tal si no disponemos del concepto lingüístico adecuado. Un estudio empírico de la relación entre pensamiento y lenguaje podría ordenar en la forma de una escala las distintas categorías de pensamientos, en función de su dependencia respecto del lenguaje. Esa escala mostraría que sin nuestra competencia lingüística sería imposible encontrarse en algunas clases de estados ϕ , o tener conciencia clara de algunas otras clases.

Tugendhat: comparten con éste la convicción de que la autoconciencia no debe interpretarse como una “percepción interna” de sí, pero rechazan que pueda consistir en un saber proposicional de los propios estados mentales, puesto que este saber proposicional mantiene la distinción entre la conciencia intencional y su objeto (la proposición ϕ), y es precisamente esta distinción la que, según Henrich, sería necesario suprimir. Ahora bien, no parece que la Escuela de Heidelberg haya propuesto hasta ahora una descripción alternativa enteramente convincente. Su caracterización positiva intenta en vano superar el escollo conceptual de una relación de conocimiento entre dos *relata* que simultáneamente deben pensarse idénticos. Este escollo plantea un dilema. Tan pronto como se distingue en el interior de la autoconciencia un polo subjetivo y un polo objetivo, surge de nuevo (al igual que sucediera en el modelo reflexivo) la dificultad de explicar la identidad entre ambos sin darla ya siempre por supuesta. Y si, para evitar esta dificultad, se opta por eliminar esa distinción, parece perderse el fenómeno mismo requerido de explicación, esto es, la conciencia *de sí mismo*.

Consideremos la primera de estas dos alternativas. Henrich quiere mantener la relacionalidad de la autoconciencia y le atribuye “una complejidad interna, una pluralidad interna”, que incluye “la diferencia entre la aprehensión y lo aprehendido”.⁴⁸ Pero como esto apenas difiere de la fatídica distinción entre un polo subjetivo y un polo objetivo en el interior de la autoconciencia, se procura mitigar la relacionalidad de la autoconciencia mediante una nueva terminología. La autoconciencia debería describirse más bien como una “familiaridad”, término éste deliberadamente escogido para sustraerse a los problemas de la reflexión, como admite expresamente Frank: “a fin de evitar los conocidos círculos en la descripción, se ha eliminado de esta expresión el pronombre reflexivo”.⁴⁹ Ahora bien, es dudoso que un simple cambio terminológico sea suficiente para resolver un problema teórico. Más bien se diría que aquí el problema queda encubierto, más que resuelto, y esto se hace evidente tan pronto como se analiza el término escogido un poco más de cerca: es obvio que esa “familiaridad” sólo puede ser familiaridad *con uno mismo*, de modo que basta con pensar hasta el final lo que está implícito en este término para que reaparezca el pronombre oportunamente eliminado y, junto con él, todos los problemas de la teoría tradicional de la autoconciencia.

Pero no es mucho más convincente la segunda opción explorada por la Escuela de Heidelberg, consistente en difuminar el momento relacional del concepto de “familiaridad”. A esta solución apuntan los pasajes en que Henrich insiste en que la “familiaridad” (con uno mismo) es *previa* a toda tematización, reflexión o conoci-

⁴⁸ Henrich, D. cit. en Frank, M.: “Fragmente einer Geschichte der Selbstbewusstseins-Theorie von Kant bis Sartre”, en M. Frank (ed.), *Selbstbewusstseinstheorien von Fichte bis Sartre*, op. cit., p. 591.

⁴⁹ Frank, M.: “Fragmente einer Geschichte der Selbstbewusstseins-Theorie von Kant bis Sartre”, op. cit., p. 592.

miento de sí (proposicional o no); o en que esa familiaridad es *inmediata*, es decir, está completamente exenta de toda mediación y de toda relación de conocimiento entre un sujeto y un objeto. El atractivo de esta solución es fácil de comprender: como parece imposible afirmar simultáneamente la identidad y la relacionalidad de los momentos de la autoconciencia, Henrich escoge la primera a costa de la segunda. Pero Henrich lleva tan lejos la eliminación de toda mediación, que su teoría conduce finalmente a un paradójico abandono de la subjetividad, y a interpretar la autoconciencia como un “proceso enteramente objetivo”.⁵⁰ Con razón critica Tugendhat esta solución, con la que Henrich y algunos de sus discípulos asimilan inadvertidamente la autoconciencia a una conciencia sin sujeto, a la mera vida consciente para la que el mundo “está de pronto ahí”.⁵¹ Y en efecto, difícilmente puede considerarse correcta una descripción de la *auto*-conciencia que elimina del fenómeno la referencia a sí mismo que precisamente lo define.

Es significativo que dos intérpretes tan distintos como Tugendhat y Frank lleguen a esta misma conclusión: el planteamiento de Henrich parece renunciar finalmente al fenómeno mismo que se proponía explicar.⁵² El propio Henrich parece percatarse de esto, y quizás por ello en algunos escritos subraya más bien el carácter enigmático, inanalizable en última instancia, de la autoconciencia:

Las circularidades en las que [la filosofía] cae en el entendimiento acerca del saber de sí terminan enseñándole que, necesariamente, ha de fracasar en su intento de dilucidar esta forma de saber del modo en que esto sucede siempre que se aprehende, en vulgar análisis, algo complejo a partir de algo más simple. (...) No podemos describirla [la autoconciencia], simplemente ya por su constitución, más que por aproximación.⁵³

Hay al menos dos razones por las que esta conclusión no puede satisfacerarnos. Es verdad que declarar inanalizable e inexplicable un fenómeno puede ser el resultado inevitable de una investigación que ha examinado hasta el final su objeto, pero también puede ser la confesión del fracaso de un enfoque teórico. En el caso de la teoría de la autoconciencia de la Escuela de Heidelberg, esto último parece ser el caso. Prescindir de una clarificación ulterior del fenómeno de la autoconciencia no parece un resultado, sino más bien una renuncia. Pero incluso si fuese cierto que la

⁵⁰ La expresión es de Pothast, un discípulo de Henrich citado por Tugendhat, E.: *Selbstbewusstsein und Selbstbestimmung*, op. cit., p. 66.

⁵¹ Henrich, D.: “Selbstbewusstsein: kritische Einleitung in eine Theorie”, en R. Bubner et al. (eds.), *Hermeneutik und Dialektik*, vol. 1, Tübingen, Mohr, 1970, p. 260.

⁵² Tugendhat, E.: *Selbstbewusstsein und Selbstbestimmung*, op. cit., p. 66; M. Frank, “Fragmente einer Geschichte der Selbstbewusstseins-Theorie von Kant bis Sartre”, op. cit., p. 598: “Lo que este modelo [el de Henrich] sigue sin aclarar es la unidad de la autoconciencia como un fenómeno a su vez consciente.”

⁵³ Henrich, D.: *Vida consciente*, Madrid, Síntesis, 2005, pp. 86, 88.

autoconciencia consiste en una “familiaridad” que no admite ulteriores análisis, una teoría completa de la autoconciencia aún debería abordar la relación entre este estadio fundamental y las formas más complejas, entre las cuales estaría la autoconciencia mediada lingüísticamente que prioriza la teoría de Tugendhat.

VI

El resultado de la controversia entre Tugendhat y la Escuela de Heidelberg es un extraño empate. En efecto, no deja de ser paradójico el hecho de que ambas teorías, al intentar evitar las dificultades de la teoría de la reflexión, terminen difuminando de uno u otro modo la especificidad de la autoconciencia. El enfoque lingüístico de Tugendhat reduce toda autoconciencia al saber proposicional acerca de nuestros estados mentales, y de este modo oscurece el fenómeno que debería explicar, al asimilar la perspectiva de primera persona a la de tercera persona e ignorar las formas más básicas de autoconciencia, sin las que resulta inexplicable el empleo significativo de los términos lingüísticos que nos permiten llevar a cabo auto-atracciones de propiedades mentales, en particular el empleo significativo del pronombre “yo”. Y por otra parte, la interpretación de Henrich de la autoconciencia como una “familiaridad” no mediada por ninguna relación de conocimiento ni de identificación del sujeto consigo mismo borra las diferencias entre la autoconciencia y la conciencia, o incluso entre la autoconciencia y los procesos objetivos no conscientes, y deja sin aclarar la relación entre ese estrato originario de autoconciencia y los niveles más complejos.

Tanto la teoría de Tugendhat como la de Henrich o Frank necesitan, por tanto, un desarrollo ulterior. No es éste el lugar para exponer ese desarrollo, por lo que tendremos que conformarnos con indicar brevemente las tareas que debe proponerse una teoría de la autoconciencia que quiera resolver las lagunas de los dos enfoques estudiados. En primer lugar, es imprescindible distinguir varios *niveles* de autoconciencia, correspondientes además a distintos estadios evolutivos tanto ontogénica como filogénicamente. Esta diferencia de niveles no queda suficientemente tematizada ni en la teoría de Tugendhat ni en la de Henrich,⁵⁴ y sin embargo es esencial para poder atribuir alguna forma de autoconciencia a los animales y a los niños pequeños y al mismo tiempo distinguirla de la autoconciencia de los sujetos humanos adultos, de la que sin duda dependen importantes propiedades como la identidad personal, la libertad o la responsabilidad moral.⁵⁵ Y en segundo lugar, una vez

⁵⁴ Manfred Frank, en cambio, ha intentado hacerse cargo de la diferencia entre los distintos niveles de autoconciencia. Cf. por ejemplo Frank, M.: “Subjektivität und Individualität”, en *Selbstbewusstsein und Selbsterkenntnis*, op. cit., pp. 9-50.

⁵⁵ Sobre la relación de estas propiedades con el lenguaje, cf. la primera parte de Tugendhat, E.: *Egozentrität und Mystik*, München, Beck, 1997.

diferenciados esos niveles, sería imprescindible analizar el modo en que los más complejos surgen a partir de los más simples, lo que probablemente sólo puede hacerse analizando las relaciones intersubjetivas y la adquisición del lenguaje.⁵⁶

Una teoría completa de la autoconciencia debería resolver las unilateralidades complementarias de Tugendhat y de la Escuela de Heidelberg. El giro lingüístico no puede llevarnos tan lejos como para admitir que la autoconciencia lingüística es la *única* forma de saber de sí de que disponemos, pero la teoría que se proponga explorar las formas que completan esta autoconciencia lingüística no puede limitarse a señalarnos una inanalizable y enigmática “familiaridad” con uno mismo.

Referencias bibliográficas

- APEL, K.-O.: “¿Es posible actualmente un paradigma postmetafísico de filosofía primera?”, en *Semiótica trascendental y filosofía primera*, Madrid, Síntesis, 2002.
- BAKER, L. R.: “The First-Person Perspective: A Test for Naturalism”, *American Philosophical Quarterly*, vol. 35, nº 4, 1998, pp. 327-348.

⁵⁶ Existen enfoques alternativos en el análisis de la autoconciencia que se proponen abordar precisamente estas tareas. D. Zahavi (*Self-Awareness and Alterity*, op. cit.) recurre a la fenomenología de Husserl para superar las aporías tanto de la teoría lingüística como de la Escuela de Heidelberg. Cf. también Tomasello, M.: “On the interpersonal Origins of Self-Concept”, en U. Neisser (ed.), *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge*, Cambridge: Cambridge University Press, 1993; Baker, L. R.: “The First-Person Perspective: A Test for Naturalism”, *American Philosophical Quarterly*, vol. 35, nº 4, 1998, pp. 327-348; Zahavi, D.: “First-person thoughts and embodied self-awareness: Some reflections on the relation between recent analytical philosophy and phenomenology”, *Phenomenology and the Cognitive Sciences*, 1, 2002, pp. 7-26; Morin, A.: “Levels of consciousness and self-awareness: A comparison and integration of various neurocognitive views”, *Consciousness and Cognition*, 15, 2006, pp. 358-371. Especialmente interesante es, en mi opinión, el planteamiento de Bermúdez, J. L.: *The Paradox of Self-Consciousness*, Cambridge, MIT, 1996; Bermúdez, J. L.: “Sources of Self-Consciousness”, *Proceedings of the Aristotelian Society*, 102, 2002, pp. 87-107. Con entera independencia del debate alemán que hemos estudiado en estas páginas, Bermúdez desarrolla una crítica a las teorías lingüísticas de la autoconciencia (presentes también en el contexto de la filosofía analítica anglosajona al que pertenece el autor) sorprendentemente similar a la de la Escuela de Heidelberg, pero su análisis de la autoconciencia prelingüística, y de su relación con las formas superiores y más complejas, se beneficia decisivamente de las aportaciones de la psicología empírica. Bermúdez muestra que el nivel más básico de la autoconciencia tiene una raíz somática y que los niveles superiores se originan en la interacción humana, que en las primeras etapas no es todavía una interacción lingüística, aunque después incorpore el lenguaje. Tanto por su punto de partida próximo a las tesis de la Escuela de Heidelberg, como por su capacidad para mostrar la emergencia de las formas más complejas de autoconciencia a partir de las relaciones de interacción y comunicación, probablemente el enfoque de Bermúdez permite resolver las unilateralidades de las dos teorías estudiadas en estas páginas.

- BERMÚDEZ, J. L.: *The Paradox of Self-Consciousness*, Cambridge (Mass.), MIT, 1996.
- BERMÚDEZ, J. L.: “Sources of Self-Consciousness”, *Proceedings of the Aristotelian Society*, 102, 2002, pp. 87-107.
- DUMMETT, M.: *Origins of Analytical Philosophy*, Cambridge (Mass.), Harvard University Press, 1996.
- FICHTE, J. G.: *Versuch einer neuen Darstellung der Wissenschaftslehre*, en M. Frank (ed.), *Selbstbewusstseinstheorien von Fichte bis Sartre*, Frankfurt, Suhrkamp, 1991.
- FRANK, M.: “Fragmente einer Geschichte der Selbstbewusstseins-Theorie von Kant bis Sartre”, en M. Frank (ed.), *Selbstbewusstseinstheorien von Fichte bis Sartre*, Frankfurt, Suhrkamp, 1991.
- FRANK, M.: *Selbstbewusstsein und Selbsterkenntnis*, Stuttgart, Reclam, 1991.
- FRANK, M.: *La piedra de toque de la individualidad*, Barcelona, Herder, 1995.
- HENRICH, D.: *Fichtes ursprüngliche Einsicht*, Frankfurt, Klostermann, 1967.
- HENRICH, D.: “Selbstbewusstsein: kritische Einleitung in eine Theorie”, en R. Bubner et al. (eds.), *Hermeneutik und Dialektik*, vol. 1, Tübingen, Mohr, 1970.
- HENRICH, D.: “Fichtes «Ich»”, en *Selbstverhältnisse*, Stuttgart, Reclam, 1982.
- HENRICH, D.: “Noch einmal in Zirkeln: eine Kritik von Ernst Tugendhats semantischer Erklärung von Selbstbewusstsein” en C. Bellut et al. (eds.), *Mensch und Moderne*, Würzburg, Königshausen & Neumann, 1989.
- HENRICH, D.: *Vida consciente*, Madrid, Síntesis, 2005.
- LAILACH-HENNRICH, A.: *Ich und die Anderen*, Berlin, Gruyter, 2011.
- LOCKE, J.: *Ensayo sobre el entendimiento humano*, México, FCE, 1994.
- MORIN, A.: “Levels of consciousness and self-awareness: A comparison and integration of various neurocognitive views”, *Consciousness and Cognition*, 15, 2006, pp. 358-371.
- SHOEMAKER, S.: “Personal Identity”, en S. Shoemaker / R. Swinburne, *Personal Identity*, Oxford, Blackwell, 1984.
- SHOEMAKER, S.: “Selbstbezug und Selbstbewusstsein”, en M. Frank (ed.), *Analytische Theorien des Selbstbewusstseins*, Frankfurt, Suhrkamp, 1994.
- STRAWSON, P. F.: *Individuos*, Madrid, Taurus, 1989.
- TOMASELLO, M.: “On the interpersonal Origins of Self-Concept”, en U. Neisser (ed.), *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge*, Cambridge: Cambridge University Press, 1993.
- TUGENDHAT, E.: *Vorlesungen zur Einführung in die sprachanalytische Philosophie*, Frankfurt, Suhrkamp, 1976.
- TUGENDHAT, E.: *Selbstbewusstsein und Selbstbestimmung*, Frankfurt, Suhrkamp, 1979.
- TUGENDHAT, E.: *Egozentrität und Mystik*, München, Beck, 1997.

- TUGENDHAT, E.: “Über Selbstbewusstsein: Einige Missverständnisse”, en Th. Grundmann *et al.* (eds.), *Anatomie der Subjektivität*, Frankfurt, Suhrkamp, 2005.
- WITTGENSTEIN, L.: *Investigaciones filosóficas*, Barcelona, Crítica, 1988.
- ZAHAVI, D.: *Self-Awareness and Alterity*, Evanston, Northwestern UP, 1999.
- ZAHAVI, D.: “First-person thoughts and embodied self-awareness: Some reflections on the relation between recent analytical philosophy and phenomenology”, *Phenomenology and the Cognitive Sciences*, 1, 2002, pp. 7-26.

José Luis López de Lizaga
Departamento de Filosofía
Facultad de Filosofía y Letras
Universidad de Zaragoza
lizaga@unizar.es