

CONCORDANCIA Y ERRORES (TIPO I Y II) EN EL DIAGNÓSTICO LOGOPÉDICO PRIMARIO¹

Antonio Fernández Cano
Mariano Machuca Aceituno
Universidad de Granada

RESUMEN

Este estudio muestra la alta concordancia entre el diagnóstico logopédico primario (test de detección) realizado por maestros/as con el diagnóstico emitido por expertos en base a la utilización de un instrumento de medida ad hoc (test diagnóstico).

Adicionalmente, se exponen los bajos errores (tipo I o II) cometidos al utilizar tal test de detección, se explora la validez del instrumento (métrica y diagnóstica) y se indican qué trastornos específicos (variables discriminantes) permiten detectar a sujetos con déficits logopédicos.

ABSTRACT

This study shows the high accordance between the primary logopedical diagnosis (detection test) made by a teacher and a technical diagnosis issued by experts using an ad hoc instrument of measurement (diagnosis test).

Additionally, it exposes the low rates of error (type I and II) derived from the use of such detection test. The validity of this instrument (metric and diagnostical) is explored. Finally it also points out which specific anomalies (discriminant variables) allow us to detect subjects with logopedical deficits.

1 Con la colaboración de Encarnación Guerrero Roldán, maestra-logopeda.

INTRODUCCIÓN

La normalidad logopédica es una cuestión harto controvertida y su diagnóstico podría ser menos fácil de lo que a *prima facie* pareciera. Se viene admitiendo, sin un fuerte fundamento conceptual, que algunos escolares presentan síntomas de trastornos en el habla, que esos síntomas son categorizables y que, en última instancia, el sujeto con un repertorio verbal correcto es distinguible del que no lo tiene. Así, los maestros y maestras de enseñanza infantil y primaria emiten diagnósticos primarios, basándose en la observación consuetudinaria, llevados de su intuición personal, sobre el estadio/capacidad logopédica de sus alumnos.

En el contexto andaluz, son los docentes quienes primero diagnostican de un modo global y genérico un posible trastorno logopédico en sus alumnos. Esos mismo alumnos serán posteriormente observados, diagnosticados y tratados, si procede, por un equipo logopédico-pedagógico. El diagnóstico primario (por primero, simple y proto-empírico) del maestro (primera aproximación al diagnóstico por anamnesis) se confronta con el diagnóstico *científico*, podríamos decir, del experto, produciéndose una situación similar a la toma de decisión en un contraste de hipótesis.

Tabla 1

TABLA DE CONTINGENCIA ENTRE DIAGNÓSTICOS (PRIMARIO vs. EXPERTOS)

	DIAGNÓSTICO PRIMARIO		
DIAGNÓSTICO	Normal	Normal	Anómalo
		Concordancia	Error Tipo I
EXPERTOS	Anómalo	Error Tipo II	Concordancia

Habría que esperar una alta concordancia entre sendos diagnósticos, aunque los errores a cometer producto de la discordancia, también son posibles.

El problema que estamos considerando no es nada nuevo, en Psiquiatría tiene una larga tradición y amplia consideración cuando se cotejan diagnósticos personales («me siento mal, doctor,») o sociales («X está necesitado de asistencia psiquiátrica») frente a diagnósticos profesionales o de expertos. En este sentido es antológico, el experimento de Rosenham (1973). Este autor y sus colaboradores simulaban tener trastornos psiquiátricos («oían voces»), solicitando asistencia médica. En la mayoría de los casos fueron diagnosticados como «insanos» y necesitados de tratamiento específico. Es evidente, en este caso, la manifestación de un alto error tipo I, al afirmar la existencia de diferencias o anomalías cuando en verdad no las hay, pues lo sujetos estaban palmaria-mente normales.

Dentro de las ciencias médicas, la concordancia en el diagnóstico es una cuestión capital dado que, en multitud de ocasiones, no existen pruebas diagnósticas que determinen con certeza absoluta la presencia o ausencia de enfermedad proporcionando un diagnóstico definitivo. Pero aunque suene paradójico, la validez de una metodología en el diagnóstico se establece en relación a un método de referencia, a partir de criterios básicos de validez: sensibilidad o capacidad para detectar a los individuos enfermos, evitando la presencia de falsos negativos, y especificidad, o capacidad para descartar a los exentos de la enfermedad, evitando la presencia de falsos positivos. Para una valoración de pruebas diagnóstico puede localizarse en Rodríguez Tuñas y Martínez de la Concha (1989). Y aunque esta cuestión no es el interés primordial del presente trabajo, sí será sometida a cierta consideración.

Nuestros médicos de atención primaria cometen en abundancia este tipo de error cuando remiten a sus pacientes «hipocondríacos» a especialistas para un segundo y más profesional diagnóstico; y es que en ciencias médicas, como ya es bien sabido, cometer un error tipo II (diagnosticar como sano a una persona no sana) es mucho más grave que el error tipo I. De aquí, la abundancia de este último.

La bibliografía sobre errores tipo I y II en investigación en ciencias sociales es extensa y prolija, véanse Darr (1987), Greenwald (1975) o Lykken (1968), como las realizaciones más consistentes. En Fernández Cano (1995, 79-82) se expone una amplia relación de trabajos que denotan la existencia del *sesgo pro estudios positivos*: los estudios deben poner de manifiesto resultados estadísticamente significativos (positivos por rechazar la hipótesis nula, H_0) como un prerrequisito de calidad científica.

Otra manifestación sutil de este sesgo en el diagnóstico sería la «profecía de autocumplimiento», expuesta por Rosenthal y Jacobson (1968), en el sentido de que la expectativa (primer diagnóstico) se ve corroborada con su cumplimiento (segundo diagnóstico).

De hecho, en este estudio estamos enfrentando dos métodos de escrutinio: un test de detección, realizado por el maestro (diagnóstico por anamnesis), menos seguro, menos articulado, menos costoso y más fácil de ejecutar, menos preciso y, sobre todo, que no constituye la base para un tratamiento posterior, frente a un test diagnóstico, administrado por un experto, más preciso, caro y difícil de ejecutar y que, ante todo, constituye la base para un tratamiento logopédico posterior.

II. MÉTODO

II.1. El problema

Este estudio tratará de dar respuesta a las tres cuestiones siguientes:

- ¿Qué grado de concordancia existe entre el diagnóstico logopédico primario realizado por maestros tutores y el diagnóstico profesional realizado por expertos?
- ¿Qué errores tipo I y II se comenten al realizar un diagnóstico primario (primera aproximación al diagnóstico por anamnesis)?

- ¿Qué variables permiten discriminar entre sujetos con déficits logopédicos diagnosticados primariamente?

Adicionalmente, un objetivo de este estudio, como paso intermedio para responder a las cuestiones que se plantean, será elaborar un instrumento de evaluación global del lenguaje espontáneo.

II.2. Diseño

Se trata de un estudio exploratorio descriptivo en el que se utilizarán criterios de validación y el análisis discriminante como diseño de investigación y técnica de análisis de datos, al par. La estructura esquemática del diseño sería:

$$\begin{array}{l} M: G_1 X_1 X_2 \dots X_n \longrightarrow Y_1 \\ M: G_2 X_1 X_2 \dots X_n \longrightarrow Y_2 \end{array}$$

Consta pues de dos grupos previamente emparejados (M: *Matched*) en base a una serie de variables intervinientes: edad, sexo y centro (los mismos valores para cada par de sujetos). Metodológicamente, el diseño sería una extensión ajustada del estudio de gemelos (*twin study*); véase la actualización que realiza Fischbein (1997) de este viejo método renacido de sus cenizas cual «ave fénix». La variable criterio (Y), diagnóstico primario del tutor/a, se dicotomiza en dos niveles: el nivel Y_1 , sujetos diagnosticados con anomalías y que configuran el grupo G_1 , y el nivel Y_2 , sujetos que no presentan déficits logopédicos a juicio de su tutor/a y que se encuadran en el grupo G_2 .

Las variables: $X_1 X_2 \dots X_n$ medidas en cada grupo serán las posibles variables discriminantes a dilucidar.

3. Variables

La variable criterio (Y) ya se ha definido como diagnóstico logopédico primario, mensurable a nivel dicotómico.

Las variables prodiscriminantes de interés serían una serie de parámetros relativos a los posibles trastornos lingüísticos, diversificadas según:

A. *Nivel elocutivo o de codificación fonológica-sintáctica*, manifestable en:

1. *Parafasias fonémicas*: fallos en la producción de los fonemas o en el encadenamiento de los mismos para formar palabras. Ejemplo: «beno» (bueno), «rifo» (grifo).

2. *Parafasias fonéticas*: errores en la realización de un fonema por fallo en los rasgos del mismo (*traits*). No se produce sustitución del sonido sino anomalías en los tiempos de emisión del mismo. Ejemplo: todos los casos de transformación fonética.

3. *Parafasias morfélicas* (morfológica o sintagmática): se manifiesta mediante cambios en la estructura de la palabra con independencia del significado de la misma. Ejemplo: «clavel» (claro), «caelas» (escaleras).

4. *Neologismos*: se trata de sustitución profunda en una palabra llegando a crear un término nuevo, alejado del original y extraño al idioma. Ejemplo: «cala» (guitarra), «pavesa» (abeja).

5. *Agramatismos*: supresión de flexiones gramaticales en género, número y/o tiempo verbal. Ejemplo: «un reyes» (un rey), «ayer voy» (ayer fui).

6. *Tanteos verbales*: son facilitaciones que realiza el propio sujeto para obtener una denominación completa y, a veces, correcta del vocablo. Ejemplo: «ga..ca..ca..racol» (caracol), «agri..ba..ero..jar..dinero» (jardinero).

B. *Nivel semántico*, manifestable en:

7. *Anomias*: se trata de la dificultad o imposibilidad para encontrar la palabra adecuada en la práctica lingüística oral. Se manifiesta en la definición por el uso, el empleo de perífrasis y/o circunloquios. Ejemplo: «ua..ua» (lobo), «hace le gesto de encender una cerilla» (cerilla), «para echar agua» (botijo).

8. *Parafasias verbales paradigmáticas* (semánticas): consistentes en la sustitución de una palabra por otra que está relacionada por el significado (relación en la red semántica) y afecta a toda su estructura. Ejemplo: «jamón» (queso), «gusano» (caracol).

9. *Parafasias monémicas*: cuando la transformación afecta a un monema (mínima unidad portadora de significado). Ejemplo: «sacapuntas» (sacacorchos), «secadora» (secamanos).

10. *Perseveraciones*: se trata de repeticiones de contenidos semánticos producidos anteriormente. Se refleja en la iteración de conceptos/términos ya enunciados y que el sujeto es incapaz de inhibir. Así, ante distintas imágenes que se le presentan para que las nombre, el sujeto nombra la última con los mismos términos que la primera.

11. *Latencia*: tiempo que tarda el sujeto en procesar la información que se le ofrece para que sea emitida; se mide en segundos. Cuando al sujeto se le presenta un objeto puede ocurrir que se tome un compás de espera que puede acabar en la producción o no. Este compás de espera o latencias se interpreta como una manifestación inhibitoria de la actividad nerviosa superior.

12. *Fatigabilidad*: se trata de una manifestación inhibitoria patológica de la actividad nerviosa superior registrable mediante la aparición de bostezos, distractibilidad, agresividad, hiperactividad motora y/o logorrea. Se mide a nivel nominal-dicotómico [1: No presencia/ 2: Sí presencia].

II.4. Instrumentos

Para medir la variable criterio se ha utilizado el juicio del profesor al considerar si un alumno necesitaba ayuda para mejorar su expresión verbal. En paralelo, se le demandaba a ese mismo profesor que indicase otro alumno de la misma edad, sexo y centro o aula que, a su juicio también, no presentase trastornos en el habla.

Para medir las variables prodiscriminantes se ha utilizado un instrumento para la evaluación del lenguaje espontáneo-intencional (en adelante, IELEI) elaborado por Machuca, González y Fernández Cano (1997).

Este instrumento trata de detectar las anomalías de la expresión en aspectos parciales de la expresión-comprensión (variables prodiscriminantes), situando al sujeto competencialmente en su oportuno nivel de codificación lingüística.

No disponemos en español de instrumentos que evalúen ampliamente el desarrollo fonemático, semántico, morfosintáctico y de vocabulario. Aportes parciales para la evaluación de tal área problemática y tenidos presente en la génesis de este instrumento serían los trabajos de Bosch (1984), centrado en el desarrollo fonológico, Serra (1984), relativo a la articulación. Monfort y Suárez (1989) elaboraron un registro fonológico baremado con población española que registra las peculiaridades del habla del niño y la compara con la producción promedio del grupo de niños de su edad. Otra prueba que explora el lenguaje oral es la PLON (Prueba de Lenguaje Oral de Navarra) de Aguinaga et al. (1989), basada en los trabajos de Bloom y Lahey (1978), y centrada sobre forma, contenido y uso del lenguaje. Se trata de una prueba relativa a grupo.

Podíamos seguir enunciando instrumentos de carácter logopédico pero todos ellos tienen, a nuestro entender, profundas limitaciones en el sentido de que:

- La mayoría son producto de traducciones anglosajonas, con lo que están culturalmente sesgados ya que como es bien sabido el español y el inglés no sólo difieren en su sistema fonológico, si no también en su estructura sintáctica, semántica y pragmática.

- Son de referencia a norma-grupo, por lo que no permiten ni indican patrones de errores susceptibles para elaborar programas terapéuticos concretos.

- No suelen contemplar la información que se tiene sobre adquisición morfológica, de manera que se aplican sin distinción a todos los niños, de todas las edades, las mismas formas verbales.

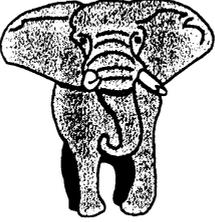
- Contemplan aspectos logopédicos parciales.

Por ello, nos pareció conveniente evaluar el estadio fonológico del lenguaje espontáneo (nivel intencional) y los procesos analítico-sintéticos inherentes al mismo a un nivel métrico. Mas, por otra parte, estimamos conveniente detectar ciertos parámetros de la comprensión que comprometen la extracción de lo significados, orientando el diagnóstico de sujetos con patología anártrica y/o afásica. En síntesis, el instrumento permite indicar el nivel de desempeño lingüístico a nivel codificativo, fonológico y semántico.

Descripción del instrumento IELEI

La prueba, tipo no verbal-verbal, trata de indagar diversos sistemas de conexiones entre el primer sistema (imágenes) y el segundo sistema de señales (lenguaje oral). En ella, el sujeto debe nombrar el objeto (imagen) que el evaluador le presenta.

La prueba consta de 150 imágenes-items, a modo de lotó, extraídas del Vocabulario Básico Curricular de Infantil y Primaria elaborado por Pérez González (1988), en la que están representados todos los elementos del código fonológico y sus combinaciones sintéticas en distintas posiciones de complejidad creciente. Para representar la prueba se sigue el criterio de desarrollo cronológico fonemático expuesto por Jakobson (1976). Véanse una imágenes-ejemplo:

		
elefante	martillo	gafas

Instrucciones para la administración

Se le informa al sujeto de la presentación de una serie de dibujos sencillos a los que tendrá que llamar por su nombre. Igualmente se le instruye a que debe anteponerle el determinante adecuado: un/una; el/la; uno/unas; los/las. Si alguna imagen no es reconocida debe manifestarlo para facilitársela bien fonológica o semánticamente.

En la hoja de respuestas se contemplan cinco apartados:

- Corrección o incorrección del vocablo emitido [SI/NO].
- Producción exacta del sujeto.
- Tiempo de latencia.
- Indicación si se oferta facilitación fonológica y/o semántica.
- Tipo de error/es si se ha/n cometido.

No existe límite de tiempo, aunque si el examinador observa fatiga por parte del sujeto se concluye la administración en una sesión posterior. El tiempo promedio ronda entre 25 y 30 minutos para alumnos de Primaria. En infantil, se completó en doble sesión de 15 minutos cada una. La administración fue individual, no existiendo límite de tiempo. La edad-suelo fue de cuatro años; la edad-techo, sin límite.

Validez y fiabilidad del IELEI

Dos aproximaciones a la validez del instrumento podríamos establecer: de constructo y de contenido. La validez de contenido estimamos que está garantizada pues queda contemplado representativamente a nuestro juicio el universo de trastornos lingüísticos admisibles (variables prodiscriminantes, anteriormente anunciadas).

Una aproximación a la validez de constructo del instrumento se ha realizado mediante análisis factorial a los componentes (trastornos/variables) contemplados. La solución más «bondadosa» e interpretable ha sido la ortogonal rotada (varimax) que manifiesta la existencia de cuatro factores específicos; ello, denota la complejidad del constructo «trastornos logopédicos» a falta de un factor general.

Tabla 2
SOLUCIÓN FACTORIAL ORTOGONAL Y ROTADA (VARIMAX) PARA COMPONENTES DEL CONSTRUCTO TRASTORNOS LOGOPÉDICOS

VARIABLE	I	II	III	IV	h ²
1. Latencia	0.89	-	-	-	0.80
2. Anomia	0.88	-	-	-	0.80
3. Neologismos	0.66	0.54	-	0.32	0.83
4. Morfémicas	0.65	0.32	-	-	0.57
5. Agramatismos	0.59	0.51	-	0.27	0.68
6. Perseveración	-	0.73	0.43	-	0.76
7. Fonéticas	-	0.70	-	-0.26	0.61
8. Parafasias V.	0.28	-	0.75	-	0.66
9. Fonémicas	0.29	-	-0.65	-	0.57
10. Tanteos	—	-	-	0.92	0.86
Varianza	2.95	1.71	1.31	1.20	

Cargas < |0.25| → 0

Una interpretación plausible de los factores obtenidos sería:

Factor 1: Cuadro afásico mixto.

Factor 2: Si excluimos la variable perseveración obtendríamos un cuadro anártrico puro y potente.

Factor 3: Cuadro con potencia afásica pura.

Factor 4: Cuadro anártrico de escasa potencia.

Obsérvese la presencia de dos trastornos genéricos bien delimitados: afásicos y anártricos. Aventuramos que una solución factorial de segundo orden aportaría evidencia más contundente a esta conjetura.

Se ha calculado una fiabilidad *interevaluadores-codificadores* del siguiente modo: Dos evaluadores independientes observan, por una parte, la presencia o ausencia de errores, y por otra, toman nota fiel de la respuesta o ausencia de la misma, que el sujeto va produciendo, relativa a todos y cada uno de los ítems presentados. Los casos discrepantes (un mismo sujeto es calificado de manera distinta ambos calificadores) son resueltos conjuntamente, volviendo de nuevo a administrar al sujeto el ítem discrepante y tomando un acuerdo final.

En una primera administración el acuerdo entre evaluadores logró un porcentaje de concordancia entorno al 86% de los casos. Los casos en que se evidenciaron discrepancia fueron 16 (14 %).

Posteriormente se calculó un índice de fiabilidad por consistencia interna de unidades, utilizando la \hat{U} de Carmines (una extensión del \hat{O} de Cronbach). El valor obte-

nido de 0,78 expresa una fiabilidad adecuada y más para un instrumento como el IELEI, inicial, exploratorio y poco depurado.

II.5. Muestra

La muestra utilizada en este estudio ha estado formada por 57 pares de sujetos de entre cuatro y doce años, de escolarización normal (en aulas ordinarias), pertenecientes a ocho centros de Jaén y provincia. Cada emparejamiento fue realizado por los propios tutores en base al mismo sexo, edad y pertenencia a misma aula o centro.

La mecánica de la selección de la muestra venía condicionada por la iniciativa de los tutores que derivaban a los alumnos para que fuesen estudiados por los Equipos de Orientación Educativa (expertos), por creer que presentaban, a consecuencia de déficits lingüísticos, dificultades en el aprendizaje. Posteriormente, se le demandó a cada profesor que presentase un sujeto «gemelo», pero que no presentase tales déficits.

La variable «sexo» está representada mostrando valores de frecuencia en torno a 68 % de varones y 32 % de mujeres. Este dato constituye de por sí un hallazgo, ya que el porcentaje de alumnos varones diagnosticados primariamente por el tutor como necesitados de atención logopédica es sensiblemente (con significación substantiva, a nuestro juicio) más alto que el de mujeres, en una proporción de 2:1.

En la variable «fetría» cabe reseñar que el 43 % de la muestra tiene al menos un hermano mayor. Este hecho es digno de destacar ya que clarifica especulaciones y puntos de vista diferentes acerca de que la alteración de la elocución suele ser afectar a niños unigénitos o que ocupan el primer lugar fetrico.

La composición de la muestra viene dada en la tabla III, teniendo en cuenta que cada valor representa un par de sujetos.

Tabla 3

DESCRIPCIÓN DE LA MUESTRA UTILIZADA EN EL PRESENTE ESTUDIO

EDAD	-os.	-as.	f_i	%
4	9	2	11	19%
5	12	6	18	32%
6	3	5	9	16%
7	6	2	8	14%
8	3	1	4	7%
9-12	5	2	7	12%
f_i	39	18	57	
%	68%	32%		100%

III. RESULTADOS

Para analizar los datos disponibles se ha utilizado la técnica multivariada del análisis discriminante (AD). Esta técnica ha estado estrechamente asociada al problema de la clasificación dado que un espacio discriminante podría ser óptimo para emitir una decisión de clasificación en ciertas situaciones (ejem. Roulon et al. 1967, pp. 299-319, 339-341). El AD sirve básicamente para aportar intuiciones sobre cómo grupos preformados (en este caso, según el criterio primario del tutor) difieren en un espacio métrico multivariado. El uso del AD en procesos de clasificación es antiguo, Bernard (1935) realizó su primera aplicación y desarrollo como técnica analítica tratando de detectar cuales eran las variaciones en cuatro series distintas de momias egipcias. Rao y Slater (1949) lo utilizaron para diferenciar entre grupos de neuróticos. Desde entonces, la técnica multivariada del AD ha tenido una fuerte implantación y desarrollo en ciencias sociales y conductuales.

III.1. Variables discriminantes

Como función informativa previa aportamos los resultados a nivel de grupo parcial (sujetos diagnosticados como anómalos frente a normales) en las variables que se consideraron prodiscriminantes a nivel de medias aritméticas.

Tabla 4

VALORES PROMEDIO DE LA PRESENCIA DE TRASTORNOS LOGOPÉDICOS
DETECTADOS POR EL IELEI

VARIABLE/ TRASTORNO	G. anómalos (A) (n = 57)	G. normales (N) (n = 57)
Fonémicas	19.2	0.0
Fonéticas	3.8	0.0
Morfémicas	2.7	0.0
Neologismos	9.3	0.0
Tanteos	2.8	0.1
Agramatismos	7.8	0.1
Paradigmáticas	9.4	17.1
Anomias	11.4	7.9
Latencias	8.1	1.5
Perseveraciones	1.1	0.0
Monémicas	0.0	0.0
Fatigabilidad	1.2	1.0

Se denotan diferencias entre grupos en la mayoría de las variables, y en la dirección plausible: más errores en el grupo A que en el N. Llama poderosamente la atención la variable «parafasias verbales *paradigmáticas*», en el sentido de que no se verifica la dirección que inicialmente intuíamos.

Aplicando el AD a los datos anteriores y tras un proceso en cuatro pasos, las variables que permiten discriminar la pertenencia de un caso/sujeto a un grupo u a otro, son las dadas en la siguiente tabla:

Tabla 5

VARIABLES DISCRIMINANTES DE LA DICOTOMIA «ANOMALIA-NORMALIDAD» LOGOPÉDICA SEGÚN DIAGNÓSTICO PRIMARIO

Paso	Variable	F	L de Wilks	F	g.l.
1	Fonémicas	96.3	0.537	96.3	1;112
2	Paradigmáticas	13.8	0.479	60.5	2;111
3	Fatigabilidad	15.3	0.419	50.7	3;110
4	Agramatismos	6.0	0.397	41.3	4;109

III.2. Concordancia entre diagnósticos

La siguiente tabla de contingencia (tabla 6), en la que aparecen los sujetos clasificados en cada grupo y el porcentaje correspondiente, nos atestigua de una fuerte concordancia entre ambos tipo de diagnóstico: el primario del tutor y el «científico» de los expertos.

Tabla 6

TABLA DE CONTINGENCIA CON CASOS (Y PORCENTAJES) DE ACUERDO Y ERRORES TIPO I Y II

DIAGNÓSTICO EXPERTOS	DIAGNÓSTICO PRIMARIO			
		Normal	Anómalo	
Normal		56 (.49) VP	8 (0.07) FP	64
	Anómalo	1 (.01) FN	49 (.43) VN	50
		57 (.50)	57 (.50)	114

En el 92 % (.49 + .43) de los casos hay concordancia entre ambos diagnósticos; incluso, adoptando un procedimiento clasificatorio más conservador cual es el del corte de navaja (*jackknife classification*), sólo aumentaría la clasificación en dos casos discrepantes, pasando de 8 a 10 los sujetos clasificados primariamente como anómalos cuando según los especialistas son normales. Tan alta concordancia nos hace tener bastante confianza en el sentido detector del tutor. De ahí la conveniencia de formular el «diagnóstico por anamnesis» en un contraste de hipótesis que el técnico del diagnóstico tendrá que confirmar o rechazar (véase Machuca, 1996).

III.3. Parámetros en la concordancia diagnóstica

Utilizando conceptualizaciones de la epidemiología, disciplina médica preocupada por cuestiones diagnósticas (i.e. Ahlbom, 1987; Janicek y Cloeroux, 1987; Rothman, 1987), podríamos obtener ciertos parámetros afines de especial relevancia. A partir de la tabla anterior y sabiendo que los casos anteriores configuran la tipología: VP (verdaderos positivos), VN (verdaderos negativos), FP (falsos positivos), FN (falsos negativos), calcularemos:

-Sensibilidad [S] del test de detección (diagnóstico primario): $S = VP / (VP + FN) = 0.89$.

-Especificidad [E] de dicho test: $E = VN / (FP + VN) = 0.87$.

El test de detección del tutor es más sensible que específico, o sea el maestro/a detecta a la mayoría de niños/as con anomalías logopédicas, pero también suele presentar como anómalos a sujetos que no lo son.

El valor global del test de detección [VG] podría calcularse: $VG = (VP + VN) / (VP + VN + FN + FP) = 0.91$, un índice bastante alto que denota la valía de tal procedimiento. Incluso, cuando se descuenta el posible acuerdo debido al efecto del azar, utilizando la K de Cohen, se obtiene un valor de 0.84, indicador muy aceptable como expresión de concordancia.

El valor predictivo del diagnóstico primario podría determinarse tanto para el test positivo (VP^+) como para el test negativo (VP^-). Así, $VP^+ = VP / (VP + FP) = 0.86$ y $VP^- = VN / (VN + FN) = 0.98$. Al ser estos dos parámetros bastante similares, podemos afirmar que la prevalencia de la «enfermedad» (trastornos logopédicos) es media a nivel poblacional ya que es bien sabido que cuanto menor sea la prevalencia de las anomalías menor será VP^+ y mayor VP^- , y viceversa en el caso de una alta prevalencia.

III.4. Tipo de errores

A partir de los datos disponibles se constata que el error tipo I (sujetos diagnosticados primariamente como anómalos pero que tras un riguroso examen son calificados normales) excede al error tipo II (sujetos diagnosticados como normales pero que necesitarían tratamiento logopédico según se constata en el examen posterior de los especialistas).

Se corrobora, una vez más, el fenómeno que denota una mayor abundancia del error tipo I (falsas diferencias en 8 casos) frente al de tipo II (falsas igualdades en 1 caso). O sea, casi un 1 % frente al 7 %. No obstante, en este contexto, no deja de ser pre-

ocupante ese caso único de error tipo II: un sujeto con apariencia de normalidad pero necesitado de tratamiento específico.

IV. CONCLUSIONES

Los hallazgos básicos que ofrece este estudio podríamos resumirlos como sigue:

— Existe una serie de variables que permiten discriminar la existencia de anomalías logopédicas en escolares de 4 a 12 años, en base a un instrumento de medida con cierta validez y fiabilidad. Tales variables son las relativas a la presencia trastornos fonémicos, paradigmáticos y agramatismos, así como la variable fatigabilidad.

— El sentido de «detección» de anomalías logopédicas en escolares que posee el profesor tutor concuerda con un diagnóstico más ajustado que emite el especialista con el concurso de un instrumento validado.

— Existen mínimas discordancias entre el diagnóstico primario del profesor y el acometido por especialistas apoyados en un instrumento de medida. Tales discrepancias expresables como errores expone una mayor tasa de error tipo I (falsas diferencias) que error tipo II (falsas igualdades) en una razón de 7:1.

V. RECOMENDACIONES

Apelar al manido «se necesita más investigación al respecto», aunque ésta sea de réplica, es el único modo de obtener una evidencia más contundente, que no definitiva, sobre las cuestiones que aquí se indagan.

Mejoras en los posibles estudios de réplica irían en conseguir grupos criteriosales lo más «gemelizados» posible que incorporen, aún a costa de más mortalidad, mayor número de variables de emparejamiento.

Igualmente, parece conveniente indagar en los mecanismos íntimos que el tutor utiliza como test de detección para diferenciar sujetos con o sin anomalías logopédicas. La dicotomización de «caja negra», que aquí hemos utilizado como variable criterial, merecería ser explorada en profundidad.

BIBLIOGRAFÍA

- Aguinaga, G.; Armentia, M.L.; Fraile, A.; Olangua, P. y Uriz, N. (1989). *Prueba de lenguaje oral Navarra (PLON)*. Pamplona: Fondo de Publicaciones del Gobierno de Navarra (Departamento de Presidencia-Interior).
- Ahlbom, A. y Norell, S. (1987). *Fundamentos de epidemiología*. Madrid: Siglo XXI.
- Barnard, M.M. (1935). The secular variations of skull characters in four series of egyptian skulls. *Annals of Eugenics*, 6, 352-371.
- Bloom, L. y Lahey, M. (1978). *Language development and language disorders*. Nueva York: John Wiley and Sons.
- Bosch, L. (1984). El desarrollo fonológico infantil: Una prueba para su evaluación. En *Estudios de psicología infantil*, M. Siguan (Ed.). Madrid: Pirámide.

- Darr, R. (1987). Another look at Meehl, Lakatos, and the scientific practices of psychologist. *American Psychologist*, 42, 145-151.
- Fernández Cano, A. (1995). *Métodos para evaluar la investigación en Psicopedagogía*. Madrid: Síntesis.
- Fischbein, S. (1997). Twin studies. En *Educational Research, Methodology, and Measurement: An International Handbook*, (2ª ed.), J.P. Keeves (Ed.), pp. 162-167. Oxford: Pergamón.
- Jakobson, R. (1976). *Lenguaje infantil y afasia*. Barcelona: Alianza.
- Greenwald, A. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, 81, 1-20.
- Jenicek, M. y Cleroux, R. (1987). *Epidemiología. Principios, técnicas y aplicaciones*. Barcelona: Salvat.
- Lykken, D. (1968). Statistical significance in psychological research. *Psychological Bulletin*, 70, 151-159. También en *The significance test controversy*, (1970) D. Morrison y R. Henkel (Eds.), pp. 267-269. Chicago: Aldine.
- Machuca, M. (1996). Hacia un enfoque comprensivo del diagnóstico en educación. *Jaenseñanza*, 10, 95-100.
- Machuca, M. y Fernández Cano, A. (1997). Elaboración de un instrumento para la evaluación del lenguaje espontáneo-intencional. (Documento interno). Granada: Facultad de Educación.
- Monfort, M. y Juárez, A. (1989). *Registro fonológico inducido*. Madrid: CEPE.
- Pérez González, J. (1988). *Didáctica de la Ortografía*. Valladolid: Miñón.
- Rao, C.R. y Slater, P. (1949). Multivariate analysis applied to differences between neurotic groups. *British Journal of Psychology, Statistical Section*, 2(1), 17-29.
- Rodríguez Tuñas, B. y Martínez de la Concha, D. (1989). Valoración de pruebas diagnósticas. En *Curso de Metodología de Investigación en Enfermería*, pp. 227-241. Granada: ICE-Universidad de Granada.
- Rosenham, D.L. (1973). On being sane in insane places. *Science*, 179, 250-259.
- Rosenthal, R. y Jacobson, L. (1968). *Pygmalion in the classroom*. Nueva York: Holt, Rinehart & Winston.
- Rothman, K.J. (1987). *Epidemiología moderna*. Madrid: Ed. Díaz de los Santos.
- Roulon, P.J.; Tiedeman, D.U.; Tatsuoaka, M.M. y Langmuir, C.R. (1967). *Multivariate Statistics for Personnel Classification*. Nueva York: Wiley.
- Serra, M. (1984). Normas estadísticas para la población de 3 a 7 años del área metropolitana de Barcelona. *Logopedia y Fonoaudiología*, 3-4, 232-235.