Globethics Repository



Polaridad de los comentarios y consistencia interna en los informes de arbitraje de artículos de investigación [Polarity of comments and internal consistency in peer review reports on scientific research articles]

This page was generated automatically upon download from the Globethics Repository. More information on Globethics see https://www.globethics.net. Data and content policy of Globethics Repository see https://repository.globethics.net/pages/policy.

Item Type	Article
Authors	Astudillo,César;Squadrito,Karem;Varas,Germán;González, Carlos
Publisher	Centro Interdisciplinario de Estudios en Bioética, Universidad de Chile
Rights	With permission of the license/copyright holder
Download date	2026-11-19 10:41:40
Link to Item	http://hdl.handle.net/20.500.12424/160634

Acta Bioethica 2016; 22 (1): 119-128

POLARIDAD DE LOS COMENTARIOS Y CONSISTENCIA INTERNA EN LOS INFORMES DE ARBITRAJE DE ARTÍCULOS DE INVESTIGACIÓN

César Astudillo[1], Karem Squadrito[2], Germán Varas[1], Carlos González[2], Omar Sabaj[1]

Resumen: El informe de arbitraje en los procesos de revisión por pares de artículos de investigación es un género clave para explicar cómo se construye colectivamente el conocimiento científico. En estos informes los evaluadores emiten, junto con una serie de comentarios, una recomendación de publicación. El análisis de la calidad del proceso de evaluación por pares se ha realizado a partir de indicadores, como las tasas de rechazo o el grado de acuerdo entre los evaluadores. Sin embargo, aún queda pendiente una evaluación cualitativa del proceso. El objetivo de este trabajo fue describir la proporción, según su polaridad (positiva, negativa y neutra), de los comentarios de 56 informes de evaluación de la Revista *Onomázein* y determinar si esa proporción era consistente con la recomendación de los evaluadores (Aceptado, Aceptado con enmiendas mayores o menores, y Rechazado). Del análisis de 1.472 comentarios se determinó que, independientemente de la decisión, la mayor proporción corresponde a comentarios negativos. Asimismo, podemos afirmar que los procesos analizados presentan un alto grado de consistencia. Mientras más favorable es la recomendación de los árbitros mayor es la proporción de comentarios positivos emitidos y, correspondientemente, menor es la proporción de comentarios negativos.

Palabras clave: proceso de evaluación por pares, consistencia, informe de evaluación, polaridad evaluativa

Polarity of comments and internal consistency in peer review reports on scientific research articles

Abstract: Peer review reports on scientific articles means a key genre to explain how scientific knowledge is collectively constructed. In these reports, reviewers write a recommendation for publication along with a series of comments. The quality analysis of the peer review process has been commonly conducted based on indicators, such as rejection rates and the agreement level among evaluators. However, a more qualitative investigation on the process still remains outstanding. This work aims at describing the polarity (positive, negative and neutral) of comments corresponding to 56 peer review reports belonging to the journal *Onomázein* and determining whether this proportion is consistent with the reviewers recommendation, i.e. Accepted, Accepted with major revisions, Accepted with minor revisions and Rejected). After the analysis of 1.472 comments, it was possible to determine that the highest proportion of comments is negative, independent of the decision. The analyzed processes also showed a high level of consistency. The more favorable the recommendation, the higher the proportion of positive comments, and, consequently, the less the proportion of negative comments.

Key words: peer review process, consistency, peer review report, evaluative polarity

Polaridade dos comentários e consistência interna nos informes de arbitragem de artigos de pesquisa

Resumo: O informe de arbitragem nos processos de revisão por pares de artigos de pesquisa é um gênero chave para explicar como se constrói coletivamente o conhecimento científico. Nestes informes os avaliadores emitem, junto com uma série de comentários, uma recomendação de publicação. A análise da qualidade do processo de avaliação por pares foi realizada a partir de indicadores, como as taxas de recusa ou o grau de acordo entre os avaliadores. No entanto, ainda resta pendente uma avaliação qualitativa do processo. O objetivo deste trabalho foi descrever a proporção, segundo a sua polaridade (positiva, negativa e neutra), dos comentários de 56 informes de avaliação da Revista Onomázein e determinar se essa proporção era consistente com a recomendação dos avaliadores (Aceito, Aceito com emendas maiores ou menores, e Recusado). Da análise de 1.472 comentários se determinou que, independentemente da decisão, a maior proporção corresponde a comentários negativos. Assim mesmo, podemos afirmar que os processos analisados apresentam um alto grau de consistência. Quanto mais favorável for a recomendação dos árbitros maior será a proporção de comentários positivos emitidos e, correspondentemente, menor é a proporção de comentários negativos.

Palavras-chave: processo de avaliação por pares, consistência, informe de avaliação, polaridade avaliativa

[1] Universidad de La Serena, Facultad de Humanidades, Departamento de Artes y Letras, Chile
Correspondencia: omarsabaj@userena.cl

[2] Pontificia Universidad Católica de Chile, Facultad de letras, Departamento de Lingüística, Chile

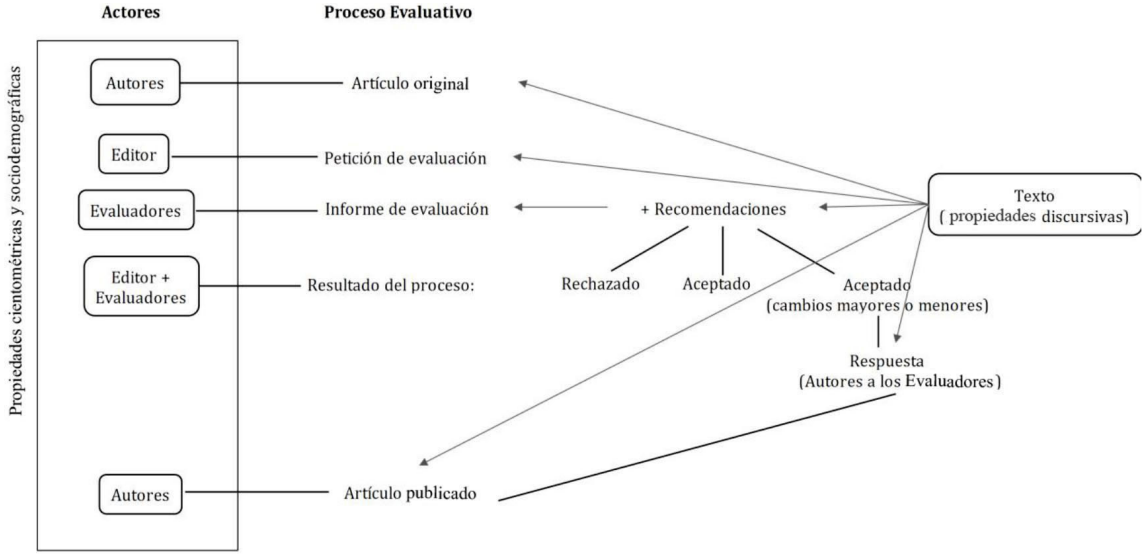
Introducción

El informe de arbitraje en el proceso de revisión por pares de artículos de investigación es un género clave para explicar la forma en que se construye colectivamente el conocimiento científico. En este proceso, el evaluador emite un informe en el que realiza una serie de comentarios y usualmente debe completar listas de cotejo o rúbricas con escalas Lickert. Asimismo, en el informe, el evaluador otorga una recomendación de publicación al editor quien, en última instancia, toma una decisión sobre la publicación del manuscrito. Dependiendo del caso, los comentarios en el informe de arbitraje pueden tener una polaridad positiva, una negativa o bien una no marcada o neutra. Entendemos por polaridad el valor que adopta el discurso cuando se usa el lenguaje con fines evaluativos(*1*). La recomendación de publicación determina el estatus de un artículo en un proceso. En la mayoría de las revistas científicas las recomendaciones se organizan en las siguientes opciones: Aceptado, Aceptado con enmiendas menores, Aceptado con enmiendas mayores, y Rechazado. Idealmente, se espera que este resultado sea consistente con la proporción de comentarios emitidos según su polaridad.

Entenderemos por "consistencia" el grado en que las recomendaciones de los árbitros se relacionan con la cantidad y proporción de comentarios negativos, positivos y neutros existentes en el informe de arbitraje. De esta manera, si un informe presenta una alta cantidad de comentarios positivos y la recomendación es aceptar el artículo, diremos que se trata de un informe consistente. De forma opuesta, si un informe presenta una alta tasa de comentarios positivos y la recomendación es rechazar el artículo, estaremos entonces frente a un informe con baja consistencia. La consistencia es muy importante desde el punto de vista de los autores del artículo, ya que, como han demostrado Bakanic y otros(*2*), estos suelen recibir mensajes contradictorios, que pueden resultar confusos, especialmente para los investigadores que se encuentran al inicio de su vida académica. Esta situación contradice lo planteado por Jefferson y otros (*3*), quienes señalan que lo deseable es que el proceso permita a los autores mejorar sus manuscritos.

El propósito de esta investigación es confirmar si existe consistencia entre los comentarios evaluativos y la recomendación de los árbitros al editor en los informes de arbitraje de una revista académica chilena.

Figura 1. El PEP como proceso sociodiscursivo complejo.



En primer lugar, presentaremos nuestra perspectiva sobre el proceso de evaluación por pares en revistas científicas (en adelante, PEP) y revisaremos antecedentes teóricos sobre el estudio de los comentarios en el PEP, específicamente en aquellos relacionados con la forma de escritura de los informes de arbitraje. En segundo lugar, describiremos las características de la revista analizada en este estudio, la muestra seleccionada y el procedimiento de análisis de los datos. Finalmente, ofreceremos los resultados, su respectiva interpretación y las principales conclusiones del estudio.

El proceso de evaluación por pares

En apariencia, el PEP es un proceso sencillo, que consta de los siguientes pasos: (a) un autor envía un artículo a una revista científica; (b) el editor selecciona evaluadores (comúnmente dos); (c) los árbitros escriben sus respectivos informes de arbitraje, en los que hacen comentarios evaluativos y entregan una recomendación al editor; (d) el editor toma una decisión, basada en las recomendaciones de los evaluadores; (e) finalmente, el editor comunica su decisión a los autores. El proceso puede contemplar más pasos, como por ejemplo que el editor pida a uno de los árbitros que evalúe los cambios aplicados a una nueva versión del texto.

Sin embargo, desde nuestra perspectiva, el PEP no es un proceso tan simple. Como puede verse en la figura 1, este puede ser concebido como un proceso sociodiscursivo complejo:

El PEP, visto desde esta perspectiva, involucra a un conjunto de actores (autores, evaluadores, editor de la revista), cada uno con propiedades sociodemográficas y cientométricas, y a un conjunto de textos que los actores intercambian entre sí durante las distintas etapas del proceso (entre estos se encuentran el artículo original, la petición de evaluación, el informe de arbitraje y el dictamen final). Cada uno de estos textos posee propiedades discursivas particulares.

Así concebido, el PEP se presenta como una interacción entre actores que poseen ciertos atributos sociales e intercambian textos con propósitos específicos: una actividad colectiva en la que los actores se coordinan epistémicamente para producir conocimiento. Esto se materializa en una práctica sociodiscursiva y sociotécnica, cuya finalidad principal es emitir un juicio o recomendación de publicación que, en definitiva, determina la producción, difusión y consumo del conocimiento científico. Todo lo anterior se lleva a cabo bajo los supuestos de la confianza e idoneidad ética de los actores involucrados, pese a los numerosos casos de fraude y críticas a este proceso*(4)*. En este trabajo nos centraremos específicamente en la naturaleza de uno de estos textos: el informe de arbitraje, texto que se caracteriza por contener una posición evaluativa, expresada en la recomendación al editor, y, generalmente, en un conjunto de observaciones que materializan comentarios específicos sobre el trabajo evaluado o partes de él.

Antecedentes

Varios estudios*(5-8)*, desde la perspectiva del análisis del discurso, han planteado que resulta extremadamente difícil estudiar los informes de arbitraje, ya que se trata de un género privado y confidencial (en términos de Swales*(6)*, un "género oculto"), por lo que su descripción ha sido muy vaga. A nuestro juicio, conocer a fondo este tipo de textos es esencial para comprender cómo se produce la ciencia de manera colectiva.

La interés por esta investigación específica surge luego de observar que la evaluación del PEP ha sido emprendida tradicionalmente a partir de la utilización de indicadores estadísticos, que poco dicen del proceso en términos de sus cualidades específicas. Estos intentos se han centrado, por lo general, en los resultados del proceso, medidos en términos de la confiabilidad, acuerdo, tasas de rechazo, y otros indicadores*(9,10)*, que implican la comparación del comportamiento de los evaluadores y los autores, sin describir ni analizar los textos que se intercambian durante el proceso. Lo que proponemos en este estudio es una forma alternativa de evaluar la consistencia interna del PEP. En la tabla 1 se presenta un resumen de algunos de los principales estudios discursivos en torno al informe de arbitraje.

Tabla 1: Resumen de las categorías de algunos estudios sobre el informe de arbitraje.

Gosden (2003)	Fortanet (2008)	Mungra y Weber (2010) Tharirian y Sadri (2013)	Bolívar (2011)
Categorías	Patrones evaluativos	I. Comentarios de contenido	Funciones discursivas
• Afirmaciones • Detalles técnicos • Referencias • Discusión • Formato	• Crítica • Recomendaciones • Preguntas • Aspectos evaluados	• Errores de ejemplo • Razonamiento científico erróneo sobre sus propios datos • Incorrecta interpretación científica de otros autores. • Errores procedimentales y falta de rigor • Irregularidades estadísticas • Falta de asociación entre aseveraciones e investigaciones anteriores • Falta de asociación entre aseveraciones y los datos • Explicación del por qué los datos son inusuales • Referencia bibliográfica incompleta • Terminología y definición	• Detractiva • Correctiva • Epistémica • Relacionante
		II. Comentarios sobre el uso del lenguaje • Comentarios sobre el léxico y sintaxis • Falta de claridad • No está bien escrito/Uso del inglés • Verborrea • Repeticiones • Error de tipeo, citación impropia o recomendación por el editor del texto • Incoherencia • Comentarios sobre el discurso y la retórica • Mejoras en el flujo de la información • Resaltar los rasgos novedosos • Atenuar o bajar el tono	

Gosden*(11)* propone un conjunto de categorías para clasificar los comentarios que suelen revisar los evaluadores, las que incluyen: (i) afirmaciones o aseveraciones, (ii) detalles técnicos, (iii) referencias, (iv) discusión y (v) formato. Fortanet*(12)*, por su parte, plantea la existencia de una serie de patrones evaluativos o tipos de acciones emprendidas por los evaluadores, las que clasifica en (i) críticas, (ii) recomendaciones y (iii) preguntas. Mungra y Weber*(13)* y Tharirian y Sadri*(14)*, por su parte, presentan una propuesta detallada para clasificar los tipos de comentarios vertidos por los evaluadores, divididos en (a) comentarios de contenido (relacionados con datos científicos), entre los que destacan la identificación de errores en el razonamiento científico de los propios datos, la falta de rigor metodológico o errores en la aplicación del procedimiento científico, y la falta de relación entre los datos y las propuestas del artículo, entre otros; y (b) comentarios relacionados con el

uso del lenguaje, los que incluyen indicaciones de errores, como falta de claridad y errores tipográficos, y recomendaciones del tipo "mejore el flujo de información". Finalmente, en el único estudio empírico en español dedicado a este tipo textual, Bolívar*(1)* propone la existencia de cuatro funciones discursivas en las que se clasifican los comentarios de los árbitros: (i) detractiva, (ii) correctiva, (iii) epistémica y (iv) relacionante.

Estos estudios, sin embargo, presentan muchas carencias que hemos revisado en detalle en otro trabajo*(15)*. Para efectos de esta investigación, el problema de los estudios presentados en la tabla 1 radica en dos aspectos: primero, la definición misma de "comentario" o la falta de criterios para la descripción de la unidad de análisis y, segundo, la inexistencia de categorías sobre la polaridad (positiva, negativa, neutra) de los comentarios, o bien la aparición de una sola (la negativa).

Una de las mayores dificultades metodológicas que cualquier estudioso del discurso enfrenta es la delimitación de la unidad de análisis*(16)*. En este caso, todos los trabajos de la tabla 1 se refieren a la descripción de "comentarios", noción que solo aparece definida en dos de ellos*(11,12)*. Tanto Mungra y Weber*(13)* como Tharirian y Sadri*(14)* restringen la noción de comentarios solo a aquellos elementos textuales que implican un cambio en el texto por parte del autor y dejan fuera, por tanto, los comentarios que no suponen la interacción, por ejemplo, cuando el evaluador describe una parte del artículo. Bolívar*(1)* se focaliza directamente en la función negativa, por lo que no analiza los comentarios positivos ni los descriptivos o neutros.

Fortanet*(12)* delimita la unidad de análisis complementando la propuesta de Suárez*(17)*. Para estas autoras un acto evaluativo es cualquier unidad estructural, independiente de su configuración léxico-gramatical, que contiene tanto el (sub) aspecto sobre el cual se comenta como la acción que se recomienda*(12:29, adaptado de 17:153)*. Gosden*(11)*, por su parte, equipara la unidad de análisis a la oración, lo que a nuestro juicio es un error metodológico, en tanto en una misma oración puede aparecer más de un comentario evaluativo.

Tanto en estos dos trabajos*(11,12)* como también en el resto de aquellos de la tabla 1, los comentarios de carácter descriptivo, esto es, los que tienen una valoración neutra, son excluidos del análisis. Fortanet*(12)* reconoce este problema cuando sostiene que, por un lado, el discurso descriptivo puede tener evaluación implícita y, por otro, que la descripción no debería ser ignorada pues evita que se generen malentendidos y sirve para mitigar las críticas que aparecen típicamente en este género. En este sentido, tal como ya hemos establecido, la mayoría de los trabajos que analizan los comentarios solo se focalizan en la polaridad negativa*(1,11)*. Algunos de estos también consideran los comentarios positivos*(12)*, pero ninguno analiza los comentarios de carácter neutro.

Otro problema metodológico de estos trabajos es que combinan en sus análisis informes con distintas recomendaciones de publicación. Como una forma de superar este problema, en este artículo solo consideramos procesos en los que ambos evaluadores dieron la misma recomendación de publicación.

El estudio de los comentarios en los informes de arbitraje

El análisis de los comentarios en los informes de evaluación es una ventana para comprender la forma en que la ciencia se construye de forma colectiva*(18)*. Bornman, Weymuth y Daniel*(19)* afirman que el tipo de comentarios realizados por un evaluador depende de las características de la revista para la que esté haciendo la revisión. De esta manera, las críticas son más fuertes en revistas de alto impacto que en las de menor impacto. Los autores señalan además que cada revista pone énfasis en distintos aspectos, razón por la cual proponen una serie de categorías temáticas, como la relevancia del estudio y el diseño de la investigación, para organizar los comentarios presentes en los informes de arbitraje.

Un estudio realizado por Bakanic, McPhail y Simon*(2)* propone que la calidad de un artículo puede medirse de mejor manera mediante la observación del acuerdo o la coincidencia que se da entre los comentarios realizados por los evaluadores de un artículo. Estos autores establecen que, si bien los revisores no siempre llegan a un acuer-

do respecto de la recomendación de publicación de un artículo, sí coinciden en diferentes criterios de evaluación, por lo que el acuerdo iría más allá de una decisión basada en aceptar o rechazar un manuscrito. Bakanic y otros*(2)* señalan también la existencia de una considerable cantidad de comentarios negativos en todos los informes, incluso en aquellos cuya recomendación final es la aceptación del trabajo. La inclinación hacia lo negativo se muestra en que los comentarios positivos tratan usualmente acerca del artículo en términos muy generales y, cuando se realizan felicitaciones, estas suelen ser someras. En cambio, los comentarios negativos poseen un carácter usualmente mucho más específico*(19)*.

En relación con esto, algunos consideran que los revisores realizan mayores apreciaciones positivas cuando el artículo presenta postulados de teorías con las que se encuentran más familiarizados, por lo que el tipo de teoría utilizada produce un sesgo intelectual*(20,21)*. Bakanic, McPhail y Simon*(22)* analizan las características del lenguaje utilizado por los árbitros en sus comentarios cuando los artículos son revisados por segunda vez, es decir, luego de ser aceptados con enmiendas en una primera instancia. Según estos autores, las segundas correcciones complejizan la tarea del editor, ya que este debe idear un nuevo tratamiento administrativo para poder reevaluar el artículo.

Finalmente, Bornmann, Wolf y Daniel*(23)* estudian la variación que se produce en los comentarios de los evaluadores en dos modalidades del proceso de revisión por pares: el proceso de revisión por pares cerrado (CPR) y el abierto (OPR). Los resultados dan cuenta de que en el primero los comentarios negativos son más específicos y aparecen en mayor cantidad que en el segundo. Según los autores, el CPR se caracteriza, además, por la presencia de mayor abundancia de marcas de cortesía hacia los autores y por una escritura más cuidada por parte de los árbitros. No obstante, se ha identificado en otros estudios*(24,25)* la presencia de lenguaje abusivo en el CPR, el cual es propiciado por el carácter anónimo de los evaluadores. Desde un punto de vista ético, esto es cuestionable, en tanto se asume que el PEP debe excluir cualquier sesgo de evaluación.

Metodología

El objetivo de la investigación fue determinar si había consistencia entre la recomendación de los evaluadores y la proporción de comentarios (positivos, negativos y neutros) en sus informes de evaluación.

La revista escogida intencionadamente para este análisis es *Onomázein*, revista chilena de lingüística, filología y traducción. Se trata de una publicación bianual, fundada en 1996 y publicada de manera ininterrumpida. La revista está indizada en la Web of Science y en Scopus. Su proceso de arbitraje es doble ciego. Cada artículo es enviado a dos árbitros sin contacto entre ellos, quienes pueden dar una recomendación de Aceptado, Aceptado con enmiendas menores, Aceptado con enmiendas mayores y Rechazado. En caso de desacuerdo en las categorías mayores (aceptado y rechazado), el artículo es enviado a un tercer árbitro para dirimir.

Como una forma de delimitar la muestra, solo se incluyeron procesos que, desde el año 2008 hasta 2012, recibieron exactamente la misma recomendación por parte de los árbitros, es decir, procesos con acuerdo total en la recomendación. Los datos analizados se resumen en la tabla 2:

Tabla 2. Datos de la muestra analizada.

Resultado	Procesos	Informes de Arbitraje	Comentarios
Rechazado	9	18	318
Aceptado con enmiendas mayores	7	14	503
Aceptado con enmiendas menores	11	22	609
Aceptado	1	2	42
Total	28	56	1472

En la tabla 2 puede observarse que se consideraron 9 artículos cuya recomendación (por parte de ambos árbitros) fue la de rechazo. Esto implica 18 informes de arbitraje, con un total de 318 comentarios. Para este estudio, hemos definido "comentario" como una proposición que se manifiesta en

una oración, una cláusula subordinada o en un sintagma no clausular, y que contiene rasgos lingüísticos evaluativos. Se consideran también comentarios (de tipo neutro) las secuencias textuales de carácter descriptivo, en las cuales no existe marca de valoración.

El total de procesos analizados fue de 28, el total de informes 56 y la cantidad de comentarios 1.472. Cada informe fue examinado separadamente por dos investigadores mediante el software *Atlas.ti 7.1*. Los textos fueron indagados en sus comentarios constituyentes, los que fueron clasificados como positivos, negativos o neutros, según correspondiera. Luego de esto, los análisis fueron confrontados para resolver las discrepancias en la clasificación. Para cuidar los datos confidenciales del proceso y cumplir con las normas éticas, los informes fueron enmascarados, de forma que ninguno de los analistas pudiese relacionar el informe con la persona que lo emitió.

Ofrecemos a continuación algunos ejemplos de comentarios extraídos de la muestra analizada:

Positivos

- Se rescata lo que pretende indagar con el trabajo
- Es un trabajo apto para ser publicado, interesante al combinar aspectos literarios y cinematográficos.
- Es un resumen muy apropiado respecto al texto que sigue.

Negativos

- No hay títulos de subsecciones.
- La utilización de ciertos términos no es adecuada para un artículo científico
- Habría que reescribir la metodología, pues no se dan datos necesarios para comprender la investigación.

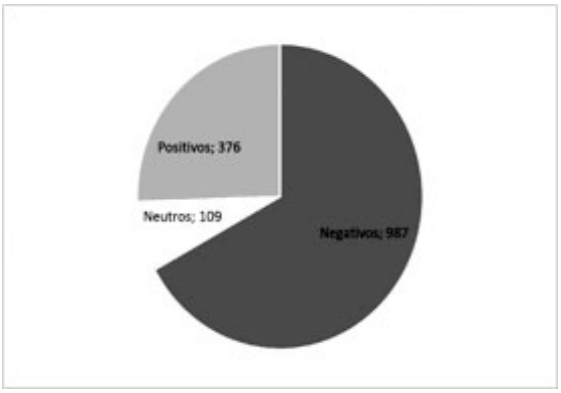
Neutros

- En el trabajo se analizan los segmentos vocálicos y consonánticos del español.
- En la metodología se muestran los procedimientos y los datos a analizar.
- Se comenta brevemente un acercamiento teórico para lenguas como el inglés.

Resultados

En el gráfico 1 puede observarse la frecuencia de los comentarios analizados, según su polaridad.

Gráfico 1. Cantidad total de comentarios por tipo en la muestra.



Un primer aspecto que llama la atención es la alta proporción de comentarios negativos (987/1.472), que alcanza casi las tres cuartas partes del total de comentarios. Esto confirma los planteamientos de Bakanic, McPhail y Simon*(2)* y Bornmann, Wolf y Daniel*(23)*, lo cual nos permite afirmar que el papel que suele asumir el evaluador de un artículo científico es el de un corrector que señala los errores del texto que evalúa. De igual forma, este resultado coincide con los de Bolívar*(1)*, quien señala que los evaluadores tienden a realizar mayormente evaluaciones de tipo correctivo, es decir, los evaluadores se focalizan en señalar las fallas en la investigación. También es llamativa la poca cantidad de comentarios neutrales (109/1.472), lo que demuestra que los árbitros no emplean su tiempo en hacer una descripción del artículo y/o sus secciones, sino que responden a la expectativa del género textual por medio de la expresión de juicios, usualmente negativos.

En la tabla 3 ofrecemos los resultados porcentuales de los tipos de comentario en relación con el tipo de recomendación expresada en el informe. En ella puede observarse que el porcentaje de comentarios negativos es consistentemente mayor

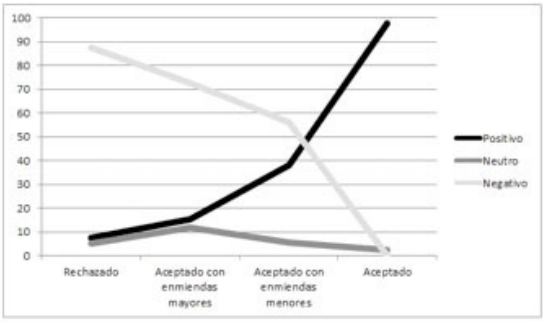
en todos los procesos analizados, independientemente del tipo de recomendación al editor, con la sola excepción de aquellos artículos en que ambos árbitros recomiendan la publicación del artículo en su estado actual, en el que solo aparecen comentarios neutros y positivos, pero ninguno negativo. En nuestros datos, la tasa de comentarios negativos por positivos es de 12 a 1 en el caso de los artículos rechazados; de 5 a 1, en aquellos aceptados con enmiendas mayores, y de 1 a 1 en el caso de los trabajos aceptados con enmiendas menores. Estas diferencias en la tasa de comentarios negativos por positivos es mucho más marcada que en otros trabajos. Por ejemplo, en Bakanic et al.*(2)* la tasa es de 5 a 1 en el caso de los rechazados y de 4 a 1 en caso de los artículos aceptados. Bornmann y otros*(19)* reportan una tasa de 6 negativos a 1 positivo en el caso de los artículos rechazados, lo que representa la mitad en comparación con los datos propios. La proporción de comentarios neutrales es siempre la más baja (con la excepción del caso recién comentado) y los comentarios positivos son también bastante escasos, salvo en los artículos cuyos informes ofrecen una recomendación de aceptado. Llama la atención la presencia de estos últimos, aunque escasos, cuando la recomendación final es "Rechazado" (7,55%). Lo anterior es un indicador de que los evaluadores no solo cumplen una función correctiva*(1)*, sino que además señalan aspectos destacables del manuscrito, aun cuando no recomienden su publicación*(24)*.

Tabla 3. Porcentaje de comentarios por tipo según la recomendación del árbitro.

	Positivo %	Neutro %	Negativo %
Rechazado	7.55	5.03	87.42
Aceptado con enmiendas mayores	15.51	11.73	72.76
Aceptado con enmiendas menores	38.26	5.42	56.32
Aceptado	97.62	2.38	0.00

El gráfico 2 muestra estos datos de manera más clara. Como puede apreciarse, la proporción de comentarios positivos asciende cuando la recomendación de los jueces es más favorable, en tanto que la proporción de comentarios negativos decrece de la misma manera. Los comentarios neutros son consistentemente bajos, pero suelen tener mayor presencia en los informes cuya recomendación es la de Aceptado con enmiendas mayores.

Gráfico 2. Tendencias de proporción de comentarios en relación con el resultado de la evaluación



La tendencia que se muestra en el gráfico 2 da cuenta de que los informes analizados son altamente consistentes, en tanto la proporción de comentarios según su polaridad (positivo, negativo y neutro) concuerda con las recomendaciones de los árbitros. De los datos analizados, podemos destacar el hecho de que solamente en los artículos aceptados la proporción de comentarios positivos es más alta que la de los negativos. En los artículos cuya recomendación es de Rechazo, la proporción de comentarios negativos es, con mucho, la más alta. De igual forma, la proporción de comentarios negativos es mayor que los positivos y los neutros en el caso de los artículos con enmiendas mayores. La explicación que proponemos a este fenómeno es simple: aunque la decisión final es positiva, los artículos en cuestión aún presentan defectos que, a juicio de los evaluadores y de acuerdo a las especificaciones formales de publicación de la revista, deben ser mejorados, por lo que estos actores hacen un esfuerzo en señalarlos con precisión. En estos artículos con enmiendas mayores es donde también priman los comentarios neutros. Dado que en estos informes los comentarios negativos son también bastante altos (72,76%), es posible que los árbitros usen los comentarios neutros como una estrategia para

introducir esos comentarios negativos. La proporción de los comentarios neutros en los artículos rechazados y aceptados con enmiendas menores es baja y muy similar, mientras que no presentan relevancia alguna en los artículos para los que se recomienda su aceptación.

Discusión

La observación conjunta de los datos nos permite afirmar que los procesos analizados en la muestra presentan un alto grado de consistencia. Así, mientras más favorable es la decisión tomada por los árbitros en la escala *rechazado>aceptado con enmiendas mayores>aceptado con enmiendas menores>aceptado*, mayor es la proporción de comentarios positivos emitidos por los árbitros y, correspondientemente, menor es la proporción de comentarios negativos.

Los resultados de esta investigación pueden ser útiles tanto para los evaluadores como para los editores. Para los primeros, pueden ser una buena fuente de información para juzgar la calidad y consistencia de sus propias evaluaciones, específicamente para que comprueben si la cantidad de comentarios según su polaridad es consistente con la recomendación entregada. Para un editor, por su parte, los datos de esta pesquisa pueden contribuir a analizar la calidad de los informes que recibe. En este sentido, si en un informe con recomendación de rechazo aparecen muchos comentarios positivos y pocos negativos, entonces, se tratará de un informe poco consistente, por lo que el editor podrá ponderar de una forma alternativa si el informe le permite o no tomar una decisión.

Agradecimientos: Este trabajo fue realizado gracias al apoyo del Proyecto Fondecyt 1130290 "La interacción socio-discursiva en la construcción colectiva del conocimiento científico: la dinámica interna del proceso de evaluación por pares". Omar Sabaj agradece además a la Dirección de Investigación de la Universidad de La Serena, por su apoyo constante.

Declaración de intereses: Uno de los autores (Carlos González) es editor de la revista desde la cual se obtuvieron los datos.

Referencias

1. Bolívar A. Funciones discursivas de la evaluación negativa en informes de arbitraje de artículos de investigación en educación. *Núcleo* 2011; 28: 59-89.
2. Bakanic V, McPhail C, Simon R. Mixed Messages: Referees' Comments on the Manuscripts They Review. *The Sociological Quarterly* 1989; 30(4): 639-654.
3. Jefferson T, Alderson P, Wager E, Davidoff F. Effects of editorial peer review: a systematic review. *Jama* 2002; *287*(21): 2784-2786.
4. Wood J. Referees and Foul Play: With scientific fraud in the News, Peer Review is once again under Attack for Missing Falsified Data. *Materials Today* 2006; 9(3): 1.
5. Bolívar A. El informe de arbitraje como género discursivo en la dinámica de la investigación. *Revista Latinoamericana de Estudios del Discurso* 2008; 8(1): 41-64.
6. Swales J. Occluded genres in the academy: The case of the submission letter. In Ventola E, Mauranen A. (eds.) *Academic writing: Intercultural and textual issues*. Amsterdam: John Benjamins; 1996: 45-58.
7. Swales J. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press; 1990.
8. Swales J. *Research genres: Explorations and applications*. Cambridge: Cambridge University Press; 2004.
9. Bornmann L, Daniel H. The validity of staff editors' initial evaluations of manuscripts: a case study of Angewandte Chemie International Edition. *Scientometrics* 2010; 85: 681–687. Doi: 10.1007/s11192-010-0215-7.
10. Cole S, Cole J, Simon G. Chance and Consensus in Peer Review. *Science, New Series* 1981; 214(4523): 881-886.
11. Gosden H. 'Why not give us the full story?': functions of referees' comments in peer reviews of scientific research papers. *Journal of English for Academic Purposes* 2003; 2: 87-10. Doi: 10.1016/S1475-1585(02)00037-1.
12. Fortanet I. Evaluative language in peer review referee reports. *Journal of English for Academic Purposes* 2008; 7: 27-37. Doi: 10.1016/j.jeap.2008.02.004.
13. Mungra Ph, Webber P. Peer Review process in medical research publication: Language and content comment. *Journal of English for Specific Purposes* 2010; 29: 43-53.
14. Tharirian MH, Sadri E. Peer Reviewers' Comments on Research Articles Submitted by Iranian Researchers. *The Journal of Teaching Language Skills (JTLS)* 2013; 5(3): 107-123.
15. Sabaj, O, González, C, Piña-Stranger, A. What we still don't know about Peer Review. *Journal of Scholarly Publishing* 2016 (en prensa).
16. Paltridge B. Genre analysis and the identification of textual boundaries. *Applied Linguistics* 1994; 15: 288-299.
17. Suárez L. *Modes of evaluation and rhetorical patterns: a contrastive study of English and Spanish book reviews*. (Tesis doctoral). España: Universidad de León; 2006.
18. Fiske D, Fogg L. But the Reviewers are making Different Criticism of my Paper! Diversity and Uniqueness in Reviewers Comments. *American Psychologist* 1990; 45: 591-598.
19. Bornmann L, Weymuth C, Daniel HD. A content analysis of referees' comments: How do comments on manuscripts rejected by a high-impact journal and later published in either a low- or high-impact journal differ? *Scientometrics* 2010; 83(2): 493-506.
20. Campanario J. El sistema de revisión por expertos (peer review): muchos problemas y pocas soluciones. *Revista Española de documentación científica* 2002; 25(3): 267-285.
21. Bornmann L, Daniel H. Reviewer and editor biases in journal peer review: an investigation of manuscript refereeing. *Research Evaluation* 2009; 18(4): 262-272. Doi: 10.3152/095820209X477520.
22. Bakanic V, McPhail C, Simon R. If at first you don't succeed: Review procedures for revised and resubmitted manuscripts. *The American Sociologist* 1990; 21(4): 373-391.
23. Bornmann L, Wolf M, Daniel HD. Closed versus open reviewing of journal manuscripts: how far do comments differ in language use? *Scientometrics* 2012; 91: 843-856.
24. Astudillo C. *Aplicación de un modelo discursivo para el análisis de los Informes de Evaluación de Artículos Rechazados (IEAR) en el Proceso de Evaluación por Pares (PEP) de tres revistas chilenas*. (Tesis de magíster). La Serena: Universidad de La Serena; 2015.
25. Campanario JM. Peer Review for Journals as it Stands Today, Part 2. *Science Communication* 1998; 19(4): 277-306.

Recibido: 27 de agosto de 2015
Aceptado: 9 de septiembre de 2015