

Variación del informe de arbitraje de artículos de investigación según la recomendación de publicación y el número de citas de los evaluadores¹

Uno de los géneros más importantes para comprender la construcción social y colectiva del conocimiento científico es el informe de arbitraje del proceso de evaluación por pares de artículos de investigación. Debido a su carácter confidencial, existen muy pocos trabajos que se dediquen a la descripción empírica de este género. Esos pocos estudios que abordan este género, a su vez, tienen una serie de limitantes, entre las que se encuentran: a) la carencia de una definición de la unidad de análisis, b) la ausencia de procesos de validación, c) la focalización en un aspecto particular del género y, d) la falta de integración de los datos discursivos con aspectos socio-relacionales. En este contexto, esta investigación tiene por objetivo mostrar los resultados de la aplicación de un modelo discursivo a un *corpus* de 318 informes de evaluación provenientes de tres revistas chilenas, agrupados según la recomendación de publicación y el número de citas de los autores.

El proceso de evaluación por pares

En términos sencillos el proceso de evaluación por pares puede caracterizarse en los siguientes pasos: a) un autor (o un grupo de autores) envía un trabajo a una revista; b) el editor recibe el manuscrito y selecciona evaluadores para que emitan un informe de arbitraje; c) teniendo en cuenta la recomendación de los evaluadores, el editor toma una decisión que informa al autor; d) Si el trabajo es rechazado el proceso concluye; por el contrario, si el trabajo es aceptado, los autores realizan las correcciones en base a las revisiones de los evaluadores y del editor y vuelven a enviar una versión para ser publicada. Este último punto puede iterarse, en la medida que el editor no quede conforme con las correcciones realizadas por el autor (Bornmann, 2011). En nuestra concepción, el Proceso de Evaluación por Pares una práctica socio-discursiva de carácter técnico y evaluativo que implica una coordinación epistémica entre distintos actores, a saber, los autores, los evaluadores y los editores (Astudillo et al., 2016). Esta práctica, que determina la generación y diseminación del conocimiento científico, es técnica y epistémica porque depende de una experticia específica de un campo de conocimiento; es evaluativa porque implica un juicio de valor por parte de los árbitros y, es discursiva, en tanto todo el proceso se cristaliza en textos que se intercambian entre los actores.

Cuatro limitaciones del Análisis del Discurso de los informes de arbitraje

Un primer y gran desafío que enfrenta cualquiera que desea analizar datos discursivos es la delimitación de la unidad de análisis (Paltridge, 1994 Swales, 2004). En el caso de los informes de evaluación, se ha utilizado como unidad de análisis la noción de comentario, sin embargo, esta unidad o no aparece definida, o bien su definición es imprecisa o parcial. Por ejemplo, en los trabajos de Mungra y Weber (2010) y de Tharirian y Sadri (2013), se limita la noción de comentarios a aquellos extractos textuales que implican una modificación por parte del autor. En segundo lugar, ninguno de los trabajos discursivos que analizan informes de arbitraje (Bolívar, 2008 y 2011; Gosden, 2003; Fortanet, 2008) describen a la vez las distintas dimensiones de esos discursos, a saber, el contenido, los propósitos comunicativos y la expresión de valoración en estos textos claves para la construcción del conocimiento científico. Otro problema de estos trabajos, que también es aplicable al campo general de los estudios del discurso, dice relación con la ausencia de procesos de validación o transferibilidad de los modelos analíticos. Cuando el análisis de datos se realiza a un nivel sintáctico o morfológico, la delimitación de la unidad de análisis y la asignación de categorías a elementos lingüísticos no necesariamente requiere de un proceso de validación de los procedimientos, pues si es el caso, se puede contar con el apoyo de informantes de esa lengua. Sin embargo, el análisis de datos discursivos es una tarea muy complicada, que depende fuertemente de los conocimientos previos del analista, por lo que validar el análisis se hace imprescindible. Es común que los trabajos en análisis del discurso no aborden el tema de la validez lingüística de las descripciones e interpretaciones que se realizan de los datos textuales. Una cuarta limitante de los trabajos dedicados al análisis del discurso de los informes de arbitraje (y también, en algunos casos del análisis del discurso en general) se refiere a la falta de información extralingüística de los sujetos que participan en la generación de estos discursos. En términos generales, el análisis del discurso utiliza en sus caracterizaciones algunos atributos sociodemográficos extralingüísticos tanto de los textos, como de los actores que los producen. En este sentido, la teoría social que subyace a la gran mayoría de los estudios del discurso de corte sociolingüístico o pragmático, realizados sea con métodos cualitativos en pequeñas muestras o con herramientas de *corpus* en grandes extensiones de texto, tiene un alcance sociodemográfico, esto es, solo da cuenta de las propiedades macrosociales de los actores participantes, como por ejemplo, su edad, sexo, inclusión en un grupo étnico, nivel o estrato socioeconómico, entre otras propiedades. A diferencia de los sociodemográficos, los atributos socio-relacionales permiten caracterizar la posición que ocupa un sujeto en relación a todos los miembros de esa comunidad, en relación a un parámetro de análisis que permite ubicar a un individuo en un *ranking*. Por ejemplo,

* Esta ponencia fue financiada gracias al proyecto Fondecyt N° 1130290, “La interacción socio-discursiva en la construcción colectiva del conocimiento científico: la dinámica interna del proceso de evaluación por pares”.

la cantidad de artículos, la cantidad de proyectos adjudicados, las citas recibidas, entre otros, son rasgos que nos permiten saber quién es quién en una comunidad disciplinar específica. La razón por la cual estos datos son relevantes se basa en la idea de que tanto las características sociodemográficas, pero sobre todo, las socio-relacionales determinan la naturaleza de los discursos de los informes de arbitraje (Varas, 2015). En este sentido se esperaría, por ejemplo, que la evaluación realizada por un investigador *senior* sea sustancialmente distinta a la de un investigador en formación. Si bien puede suceder que este no sea el caso, por la falta de esta información en los trabajos revisados, no es posible asociar las características de los participantes a los tipos de actos que realizan en este género, lo cual se constituiría en un aspecto relevante para caracterizar de forma más exhaustiva este género clave en la producción colectiva del conocimiento científico.

Método

Los objetivos de la investigación fueron dos: i) determinar la variación discursiva del informe de arbitraje (en relación a los objetos evaluados y los propósitos comunicativos) según la recomendación de publicación y, ii) analizar la variación discursiva del informe de arbitraje según el número de citas de los evaluadores

Datos

Como una forma de superar la falta de integración de los aspectos sociométricos y discursivos en el análisis del proceso de revisión por pares, se recolectó una muestra de 318 informes de evaluación, provenientes de una población de 738 informes, asociados a 369 procesos de evaluación, correspondientes al total de artículos enviados durante los años 2008 y 2012 a tres revistas internacionales publicadas en español: *Onomázein*, *Información Tecnológica* y *Formación Universitaria*. Para el desarrollo del primer objetivo se registró la recomendación de publicación del total de la muestra (318 informes). Para cumplir con el segundo objetivo, se estableció una submuestra de 118 informes, en la que se registraron el número de citas de los evaluadores de una de las revistas analizadas. Como una forma de cuidar la confidencialidad de los datos, los asistentes encargados de registrar el número de citas solo tuvieron acceso al listado de evaluadores por parte del editor, sin que fuera posible vincular qué evaluador revisó qué artículo. Para contar el número de citas, primero se determinó si el evaluador tenía un perfil de citaciones en Google Scholar. Cuando este no era el caso, se buscó a los evaluadores en otras bases de datos que contienen números de citas, a saber, Web of Science y Scopus. Por último, para los casos de los evaluadores que no aparecían en ninguna de esas bases de datos se generó un procedimiento de conteo manual, a través de la búsqueda del listado de sus publicaciones en Google Scholar, donde fue posible encontrar el número de citas para cada una de las publicaciones. La conformación de la muestra de informes de evaluación, aparece en la Tabla 1, separada por cada revista y por la recomendación de publicación de los evaluadores:

Tabla 1: Conformación de la muestra

Recomendación	Onomázein	Formación Universitaria	InformaciónTecnológica
Aceptado	17	--	--
Condicionado	61	76	58
Rechazado	40	24	42
TOTALES	118	100	100

Para la consecución del segundo objetivo, se trabajó con los 118 informes de la revista *Onomázein*. Para la obtención de los resultados de este segundo objetivo, se agruparon los datos en cuatro grupos de acuerdo a la distribución de las frecuencias de las citas al año 2014 de cada evaluador, tal como se muestra en la Tabla 2:

Tabla 2: Número de evaluadores por rango de citas recibidas

Rango citas	N
0-50	69
51-100	22
101-200	15
+200	12
TOTAL	118

Procedimientos para el análisis del discurso de los informes de arbitraje, obtención de resultados y cálculos estadísticos

En Sabaj et al. (en preparación), puede encontrarse el listado de categorías del modelo discursivo utilizado para el etiquetaje. El modelo contiene tres capas que permiten describir de forma integrada, el contenido de un texto, los propósitos comunicativos y la expresión de la valoración. Por razones de espacio, en lo que sigue solo se mostrarán los resultados relativos a dos de los componentes del modelo, a saber, el *objeto evaluado* y los *propósitos comunicativos*. Este modelo fue validado a través del cálculo del alfa de Cronbach (0,94) y el cálculo del acuerdo inter-codificador (0,65). El etiquetaje de los datos discursivos lo realizaron 14 asistentes, quienes, al igual que en el proceso de construcción del modelo discursivo de análisis trabajaron en parejas. Específicamente, los etiquetadores siguieron el siguiente procedimiento: i) Identificar el límite del comentario, ii) Determinar si se trataba de un comentario ajeno; en caso positivo, el análisis concluía en ese punto, iii) Determinar si el comentario era descriptivo o evaluativo. En el primer caso, se le asignaba el código

correspondiente y el análisis concluía en ese punto, iv) En el caso de los comentarios evaluativos, se aplicaron, primero, la categoría del contenido u objeto evaluado y, luego, el propósito comunicativo. Para la obtención de los resultados se contaron las frecuencias y se calcularon los porcentajes para cada categoría. En otro trabajo pueden encontrarse, aunque no sobre los mismos datos, la aplicación del micro-objeto (Astudillo, 2015) y del componente de la valoración (Astudillo et al., 2016). Los resultados de la aplicación del total de las categorías sobre la muestra que acá se analiza están disponibles para los interesados por medio de una solicitud a los autores.

Resultados y discusión

En la Tabla 3, se muestran los porcentajes asociados a los objetos evaluados según la recomendación de publicación.

Tabla 3: Porcentaje de objetos evaluados según la recomendación de publicación

Objeto evaluado	Aceptados	Condicionados	Rechazados	Total
Título	3	2	1	2
Resumen	6	4	5	4
Introducción	1	4	3	4
Marco teórico	1	2	1	2
Metodología	1	8	15	11
Resultados	4	8	8	8
Conclusión	5	3	3	3
Referencias	10	9	7	9
Organización interna	1	1	1	1
Artículo	68	59	56	58
Total	100	100	100	100

Tal como se muestra en la Tabla 3, independiente de la recomendación de publicación, el objeto evaluado que prima es el artículo en su totalidad. Es interesante destacar que, luego de esa categoría, la parte más relevante de los artículos aceptados son las referencias, aspecto que también es prominente en el caso de los artículos condicionados. En este tipo de informes, las categorías más relevantes son la metodología y los resultados. Lo mismo sucede en el caso de los artículos rechazados, es decir, los artículos se rechazan principalmente por problemas en esas secciones. La Tabla 4 muestra los resultados de los propósitos comunicativos según la recomendación de la publicación.

Tabla 4: Porcentaje de propósitos comunicativos según la recomendación de publicación

Propósito comunicativo	Aceptados	Condicionados	Rechazados	Total
Condiciona publicación	1	3	1	2
Corrige	10	10	7	9
Describe	3	12	15	13
Destaca error	7	19	32	24
Describe lo que falta	2	7	13	9
Destaca aspecto positivo	44	11	4	8
Evalúa aporte	9	2	3	3
Evalúa organización	1	1	1	1
Indica lo que sobra	1	1	1	1
Rechaza la publicación	0	0	2	1
Recomienda publicación	1	1	0	1
Solicita acción	5	15	8	12
Sugiere acción	10	7	4	6
Sugiere acción mediante pregunta	0	4	4	4
Sugiere inclusión	6	6	4	5
Sugiere inclusión mediante pregunta	0	2	1	1
Total	100	100	100	100

Se puede observar que algunos propósitos comunicativos son independientes de la recomendación, por lo que tienen una frecuencia relativa muy parecida en los informes según el tipo de recomendación (Corrige, Sugiere inclusión y Sugiere inclusión mediante pregunta). La variación de la recomendación de publicación en términos

discursivos se puede evidenciar en la mayor ocurrencia de ciertos propósitos para algunas de las recomendaciones. Por ejemplo, el propósito “Destaca aspecto positivo” es característico de los informes de los artículos aceptados; los propósitos “Destacar error” y “Solicitar acción” son frecuentes en los informes de artículos condicionados, y las tres categorías “Describe”, “Destaca error” y “Describe lo que falta” que son propósitos comunes de los informes de arbitraje de artículos rechazados.

Tabla 5: Porcentaje de objetos evaluados según el número de citas de los evaluadores

Objeto evaluado	0-50	51-100	101-200	más de 200	Total
Título	2	1	6	2	2
Resumen	4	6	4	2	4
Introducción	3	2	1	1	3
Marco teórico	2	6	1	4	3
Metodología	10	8	13	12	10
Resultados	6	6	6	5	6
Conclusión	3	1	3	2	2
Referencias	8	10	8	9	9
Artículo	61	59	55	61	60
Organización interna	1	1	3	1	2
Total	100	100	100	100	100

En la Tabla 5, se muestran los objetos evaluados según el número de citas de los evaluadores. Los datos sugieren que el número de citas de los evaluadores no implican grandes variaciones en términos de los objetos evaluados en los informes de arbitraje de artículos de investigación. Al igual que en el caso de las variaciones según el tipo de recomendación, es el artículo como un todo en el que los evaluadores centran sus evaluaciones. Independiente de las citas de los evaluadores aparecen los comentarios a la metodología, a los resultados y a las referencias como categorías homogéneas con una presencia muy similar. Algunas categorías muestran diferencias de uso según las citas de los evaluadores. Por ejemplo, los árbitros con un rango de citas entre 51 y 100 se focalizan más en el marco teórico, mientras que los que tienen entre 101 y 200 citas son los que más se centran en el título. En general, como se observa en los datos, son pocas las categorías que permiten diferenciar la influencia del número de citas en los objetos evaluados en los informes de arbitraje de artículos de investigación.

Tabla 6: Porcentaje de propósitos comunicativos según el número de citas de los evaluadores

Propósitos comunicativos	0-50	51-100	101-200	Más de 200	Total
Condiciona publicación	2	1	1	1	2
Corrige	10	9	7	10	9
Describe	9	12	7	11	9
Destaca error	30	27	37	27	30
Describe lo que falta	8	6	10	5	7
Destaca aspecto positivo	8	11	7	7	8
Evalúa aporte	3	2	2	1	2
Evalúa pertinencia	0	1	1	0	0
Indica lo que sobra	1	1	1	1	1
Recomienda publicación	0	1	0	0	0
Solicita acción	13	8	11	10	11
Sugiere acción	6	9	7	9	7
Sugiere acción mediante pregunta	3	1	2	8	4
Sugiere acción mediante signo de puntuación	0	1	0	0	0
Sugiere inclusión	7	11	7	10	8
Total	100	100	100	100	100

En la Tabla 6, se exponen los resultados de los porcentajes de aparición de los propósitos comunicativos según el número de citas de los evaluadores. Los datos reflejan, en términos generales, que el número de citas de los evaluadores generan variaciones en la proporción de algunos de los propósitos comunicativos de los informes de

arbitraje de artículos de investigación. Las diferencias, aunque mínimas, se dan en la proporción de propósitos específicos que permiten dar cuenta de la forma en que el número de citas de los evaluadores determinan el discurso de los informes de arbitraje. Por ejemplo, es interesante el hecho de que los evaluadores con menos citas son los que más solicitan acciones al autor y que los evaluadores que tienen entre 101 y 200 citas son los que más destacan errores y describen lo que falta. En el resto de los propósitos no se encontraron mayores variaciones según el número de citas de los evaluadores. Es muy complejo establecer las causas de estas diferencias. Varas (2015) establece los evaluadores que cumplen más funciones en una revista, esto es que participan como autores y como evaluadores en varias ocasiones tienen más citas y en sus informes tienden más a hacer sugerencias que a solicitar acciones de forma directa. Esto, según Varas (2015) se debe a que el estatus de los evaluadores con más citas hace que estos no se sientan obligados a solicitar acciones específicas a los autores, sino que asumen una posición “iluminadora” que sugiere el camino para mejorar el trabajo, pero no lo indica de manera directa. En cambio, los evaluadores con poco estatus, como una forma de validarse ante el editor, asumen una posición de poder más directa, lo que se refleja en el mayor porcentaje de solicitudes de acción directiva.

Conclusiones

Cuando se analiza un texto confidencial como el informe de arbitraje, en el cual se materializa lo que comúnmente se denomina “la construcción colectiva del conocimiento científico” es necesario incorporar algunas variables extralingüísticas que comúnmente son excluidas del análisis del discurso. A partir del análisis realizado, podemos concluir que la recomendación de publicación, esto es, si el artículo evaluado fue aceptado, condicionado o rechazado incide más en la variación discursiva del informe de arbitraje que el número de citas de los evaluadores.

Aun cuando en este estudio en particular esta última variable no genera una gran variación en los informes, es relevante incorporarla como parte de los análisis del discurso en el contexto científico. En efecto, creemos que estos rasgos extra-lingüísticos (en este caso, el número de citas, pero en términos más generales, cualquier variable continua como el número de artículos publicados, el número de instituciones en las que trabaja o el número de co-autores con los que escribe) son rasgos importantes de ser explorados en su relación con los discursos que esos actores producen.

Referencias bibliográficas

- Astudillo, C. (2015). 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. Universidad de La Serena. La Serena, Chile.
- Astudillo, C; Squadrito, K; Varas, G; González, C. y Sabaj, O. (2016). Polaridad de los comentarios y consistencia interna en los informes de arbitraje de artículos de investigación. *Acta Bioethica* 22 (1), 119-128.
- Bornmann, L. (2011). Scientific Peer Review. *Annual Review of Information Science and Technology* 45 (1), 7–245.
- Bolívar, A. (2008). El informe de arbitraje como género discursivo en la dinámica de la investigación. *Revista ALED* 8 (1) 41-64.
- Bolívar, A. (2011). Funciones discursivas de la evaluación negativa en informes de arbitraje de artículos de investigación en educación. *Núcleo* 28, 59-89.
- Fortanet, I. (2008). Evaluative Language in Peer Review Referee Reports. *Journal of English for Academic Purposes* 7, 27–37.
- Gosden, H. (2003). ‘Why not give us the full story?’: functions of referees’ comments in peer reviews of scientific research papers. *Journal of English for Academic Purposes* 2 (2) 87–101.
- Mungra, Ph. & Webber, P. (2010). Peer Review process in medical research publication: Language and content comment. *Journal of English for Specific Purposes* 29, 43-53.
- Paltridge, B. (1994). Genre analysis and the identification of textual boundaries. *Applied Linguistics* 15, 288-299.
- Sabaj, O.; González, G.; Astudillo, C.; Varas, G.; Fuentes, M.; Cabezas, P.; Squadrito, K. & Pina-Stranger, P. (en preparación). [en línea]. Anexo 1: Un modelo socio-discursivo para el análisis de los informes de evaluación en el proceso de evaluación por pares y su aplicación a los procesos de revisión de tres revistas chilenas internacionales Disponible en: <https://omarsabaj.files.wordpress.com/2007/02/anexo-1-un-modelo-para-el-ad-de-los-ia.pdf>
- Swales, J. (2004). *Research genres: Explorations and applications*. Cambridge: Cambridge University Press.
- Tharirian, M. H. & Sadri, E. (2013). Peer Reviewers’ Comments on Research Articles Submitted by Iranian Researchers. *The Journal of Teaching Language Skills* 5 (3), 107-123.

Varas (2015). El informe de arbitraje en el proceso de revisión por pares de artículos de investigación: niveles de retroalimentación según el tipo de evaluador. Tesis de Magíster. Universidad de La Serena. La Serena, Chile.