

MULTISEMIOSIS Y LINGÜÍSTICA DE CORPUS: ARTEFACTOS (MULTI)SEMIÓTICOS EN LOS TEXTOS DE SEIS DISCIPLINAS EN EL CORPUS PUCV-2010*

MULTISEMIOSIS AND CORPUS LINGUISTICS: MULTISEMIOTIC
ARTIFACTS IN THE TEXTS OF SIX DISCIPLINES IN THE
ACADEMIC PUCV-2010 CORPUS

GIOVANNI PARODI

Pontificia Universidad Católica de Valparaíso. Valparaíso, Chile
gparodi@ucv.cl

RESUMEN

Desde los estudios estrictamente lingüísticos, la caracterización de los textos escritos especializados desde su realidad multisemiótica ha sido escasa o casi nula. No son abundantes las investigaciones, basadas en corpus, que aborden la descripción de artefactos tales como fórmulas, gráficos, tablas y diagramas así como de sus disposiciones en la página escrita en textos académicos. En este contexto, el objetivo de esta investigación es identificar, describir y cuantificar la ocurrencia de los artefactos (multi)semióticos presentes en una muestra de textos (1.043) en seis disciplinas científicas del Corpus Académico PUCV-2010. El corpus fue recolectado en doce programas de doctorado en seis universidades chilenas y se compone de los materiales de lectura que leen los alumnos con la excepción de los que contempla la tesis de grado (3.160 textos escritos en las disciplinas de: Física, Química, Biotecnología, Historia, Literatura y Lingüística). Como parte de los resultados, se identifican y definen nueve artefactos multisemióticos, basados en cuatro sistemas semióticos interactivos: verbal, matemático, gráfico y tipográfico. Emergen interesantes diferencias entre los textos de las Ciencias Sociales y Humanidades y los de las Ciencias Básicas, debido a los prototipos de artefactos caracterizadores y su ocurrencia a través de las disciplinas.

Palabras clave: Multimodalidad, artefactos multisemióticos, discurso disciplinar doctoral escrito.

* Proyecto Fondecyt 1090030.

ABSTRACT

From strictly linguistic studies, the characterization of multisemiotic written specialized texts has been scarce or almost null. Not many corpus-based studies focus on the description of graphs, tables, and diagrams, as well as their layouts, as part of academic texts. The objective of this study is to identify, describe and quantify the occurrence of (multi)semiotic artifacts which are present in a sample of texts (1.043) belonging to the PUCV-2010 Academic Corpus. The corpus was collected in twelve PhD programmes in six Chilean universities and comprises all the documents students are given to read during their formal curricula, with the exception of those included in the final doctoral research (3,160 written texts, which are distributed among Physics, Chemistry, Biotechnology, History, Literature, and Linguistics). As part of the results, nine artifacts were identified and defined, based on the distinction of four semiotic interacting systems: verbal, graphical, mathematical, and typographical. Interesting differences are detected between Basic Sciences (BS) and Social Sciences and Humanities (SS&H), due to the prototypical type of artifacts and also considering the quantification of their occurrence across disciplines.

Keywords: Multimodality, multisemiotic artifacts, disciplinary doctoral written discourse.

*Recibido:*23/07/2010. *Aceptado:*14/09/2010

1. INTRODUCCIÓN

Desde los estudios estrictamente lingüísticos, la caracterización de los textos escritos especializados desde su realidad multisemiótica ha sido escasa o casi nula. No son abundantes las investigaciones, basadas en corpus, que aborden la descripción de artefactos tales como fórmulas, gráficos, tablas y diagramas así como de sus disposiciones en la página escrita en textos académicos. Así, la inherente multisemiosis de los textos escritos especializados emerge como una realidad no totalmente asumida por muchos lingüistas ni necesariamente explicada a cabalidad por sus defensores. La tendencia ha estado mayoritariamente orientada al monopolio del sistema verbal o lingüístico, muy posiblemente porque el concepto de lenguaje ha estado dominado por visiones aún muy inmanentistas y porque solo el estudio del sistema verbal impone desafíos soberbios aún no indagados completamente.

De hecho, nuestros propios estudios en los últimos años han estado enfocados al análisis de lo lingüístico *stricto sensu*, aunque desde una perspectiva multidimensional y también retórico funcional (Parodi, 2005, 2007a y b, 2010a y b). Tanto las investigaciones en torno al Corpus PUCV-2003 como al Corpus PUCV-2006 han abordado preferentemente el sistema verbal y solo han consignado de modo parcial los rasgos no verbales. Probablemente ello se explica por el tremendo afán

impuesto en desarrollar y llegar a contar con un etiquetador morfosintáctico para el español junto a una base de datos amplia y robusta y que posibilite la consulta en línea, tal como se materializó en el proyecto El Grial (www.elgrial.cl) (Parodi, 2006, 2010a).

Paralelamente a estos afanes descriptivos y tecnológicos, progresivamente más complejos debido a la cada vez más profunda mirada por parte del investigador del fenómeno en estudio, también existen otras problemáticas científicas muy ligadas a la naturaleza de los textos escritos. Estas dicen relación con la alfabetización académica y profesional (Parodi, 2010d). Como bien se sabe, desde el punto de vista de la formación universitaria, se exige a los alumnos el manejo eficiente de la lengua escrita, con el fin de que puedan comprender y escribir los textos de los diversos géneros especializados. Desde una mirada ontogenética, es posible afirmar que los seres humanos nacen más proclives hacia el modo oral de la lengua y que, progresiva y paulatinamente, se van especializando en el modo escrito; por ende, la inclinación de un estudiante por un área disciplinar conlleva el necesario aprendizaje de los sistemas semióticos especializados de los textos de su disciplina. Pero: ¿cuáles son las diferencias de los géneros de las disciplinas que un estudiante debe aprender a procesar psicolingüísticamente desde textos multisemióticamente complejos?, ¿cómo se aprende a construir significados combinando sistemas tales como el verbal, el visual, el matemático y el tipográfico? Y más específicamente: ¿cómo se enseñan unos y otros sistemas de elaboración de significados y cómo se aprende a comprenderlos cuando se leen y a escribirlos cuando se textualizan? Es muy cierto que conocemos de las diferencias entre los géneros discursivos de unas y otras disciplinas, pero exactamente de qué naturaleza son estas y qué implicancias tienen para la formación académica y profesional en ámbitos diversos son cuestiones poco o nada exploradas.

Ahora bien, en esta línea de desarrollos progresivos y acumulativos, el objetivo específico de este trabajo es llevar a cabo una identificación y descripción de los artefactos (multi)semióticos del discurso especializado en seis disciplinas científicas a partir del Corpus PUCV-2010. También buscamos comparar la ocurrencia de estos artefactos en los textos a través de las seis disciplinas, explorando algunas variables como ocurrencia promedio por página. Cabe hacer notar que en esta primera aproximación a los datos de este corpus, enfatizamos una mirada disciplinar del discurso académico escrito que los estudiantes de doce programas de doctorado deben enfrentar como lecturas durante su formación universitaria. Así, en esta fase de la investigación no establecemos distingo por género discursivo, pues el foco está puesto en el estudio del discurso académico disciplinar y su caracterización a partir de los artefactos multisemióticos que se identifican en una muestra aleatoria y proporcional del 33% del corpus total. De este modo, este estudio se centra en la interface entre cuatro sistemas semióticos de los textos escritos: el verbal, el gráfico, el matemático y el tipográfico.

La investigación mayor, en la cual este trabajo particular se inscribe, tiene como objetivo el llegar a conocer los géneros del discurso que circulan como lectura en doce programas de doctorados y en seis disciplinas particulares. También se pretende llegar a construir un perfil de competencia de la comprensión de textos especializados en las disciplinas por parte de alumnos de estos programas de doctorado por medio de metodologías experimentales y con apoyo de tecnologías computacionales. Así, para alcanzar este objetivo final es fundamental, en una primera fase, identificar las características multisemióticas del discurso académico en indagación, conocer los géneros más prototípicos y describir sus otros diversos rasgos más prominentes.

En la primera parte de este artículo entrego algunas reflexiones acerca de la multisemiosis y su carácter de rasgo fundante en los textos escritos. Luego se da cuenta del corpus y los procedimientos de recolección. Posteriormente, se detallan los aspectos metodológicos y se procede al análisis de resultados. Se cierra el artículo con algunos comentarios finales y proyecciones.

1. LA MULTISEMIOSIS: UNA BREVE PANORÁMICA

1.1. Distinciones verbal/no-verbal y paratexto

Como se sabe, clásicamente se ha aceptado la distinción entre lenguaje verbal y lenguaje no-verbal, empleando así también el término *lenguaje* para recursos de diversa índole tales como corporales, gestuales y –en general– de tipo visual. Es decir, todo aquello estrictamente no lingüístico que en conjunto a lo exclusivamente verbal dan forma a un determinado texto que pertenece a un género del discurso, como por ejemplo una conversación casual cara a cara o un artículo científico. En el caso de la lengua oral, estos recursos extra-lingüísticos fueron los primeros sobre los que se puso atención, de modo que se llegó a hablar de la comunicación mediante las manos, el tacto, la postura corporal, el olfato, etc. En un libro clásico, Davis (1976) apunta que el interés por la comunicación no verbal emerge de una tendencia a una desconfianza hacia las palabras, pues estaríamos más orientados visualmente “...más abiertos a la comunicación visible, corporal” (1976: 252). Junto a este texto, existen otros dos clásicos, uno escrito por Poyatos (1983) y otro editado por el mismo autor (Poyatos, 1992). En estos trabajos se exploran las relaciones entre lo verbal y lo no-verbal en ámbitos como la literatura, la lingüística, la sociología y la antropología, entre otros. Así, los objetivos declarados en esta última obra son: incluir áreas no abordadas previamente, impulsar la investigación en focos poco o nada explorados y conseguir mayor reconocimiento para los estudios en comunicación no-verbal.

Siguiendo esta línea de antecedentes, en el caso de los textos escritos, el *para-*

texto tradicional y etimológicamente ha sido definido como “todo lo que rodea o acompaña al texto” o “todo lo que queda de un libro o publicación cuando se separa el texto principal”. En esta línea, Alvarado (1994), siguiendo a Genette (1997), propone la existencia de tres dimensiones del paratexto: verbal, icónico y material. Para efectos de nuestra investigación, este concepto es relevante pues se pone en el foco el denominado paratexto icónico en el cual se inscribirían –entre otros- figuras, fotografías, dibujos y tablas en un texto estático y escrito. Existen otras diversas clasificaciones de paratextos, incluso algunas proponen que se debe distinguir un paratexto de tipo infográfico. Las clasificaciones en general están basadas desde el punto de vista de quien escribe el texto, esto es, se busca distinguir el autor, del editor, del diseñador, etc. Ello quiere decir que el marco de análisis del concepto de paratexto está directamente vinculado a la autoría de las secciones o agregados que se hacen. En suma, el paratexto se circunscribe a ciertas secciones o apartados de texto, a variaciones tipográficas y de diagramación de texto así como a la gráfica que acompaña a la página. Todos estos elementos, sean verbales, icónicos o infográficos, se consideran –la mayoría de las veces– en una función auxiliar y no medular, pues –desde esta perspectiva– el texto mismo en su carácter eminentemente verbal sería el núcleo (Alvarado, 1994; Genette, 1997).

Así las cosas, resulta muy relevante llamar la atención hacia esta distinción que se establece entre el texto y el paratexto, y destacar que no solo se considera así en la línea de estos autores. Al parecer, existe una cierta confusión emergente en torno al concepto mismo de texto y la decisión –de hecho no siempre evidente– de si otros sistemas semióticos se constituyen en rasgos inherentes a todo texto o pasan a conformar una cualidad adherida o de tipo satelital al texto mismo que, en este caso, quedaría circunscrito solo a lo puramente verbal o, dicho de otro modo, a lo lingüístico. Como se aprecia, a propósito de la noción de paratexto se revela un problema emergente, esto es, el concepto de texto y el de la naturaleza misma del lenguaje humano.

Desde este enfoque y los autores hasta aquí comentados, se hace evidente que, en términos globales, estos estudios han establecido una distinción y han separado el sistema verbal de otros posibles. Se rescata y valora lo no-verbal, pero no se concibe como inherente al texto mismo o se deja sentir cierta ambigüedad sobre la naturaleza multisemiótica de los textos.

1.2. El descubrimiento de lo obvio: la multisemiosis del texto escrito

No obstante los aportes comentados más arriba, concordamos con Scollon y Scollon (2009) en que la multisemiosis del texto escrito no constituye simplemente una nueva manera de mirar a un ámbito clásicamente denominado no-verbal. Es ciertamente mucho más que eso.

Ahora bien, en este apartado no se pretende ni muy someramente realizar una

revisión de los actuales estudios acerca de la multisemiosis del texto escrito. Nuestro objetivo es establecer un marco conceptual básico que encuadre nuestra investigación y así situar nuestro punto de vista en este tema que abordamos. De entrada, baste decir que partimos del supuesto fundamental de que todo texto escrito es esencialmente de naturaleza multisemiótica y que, progresivamente, pretendemos hacernos cargo de ello. Desde esta perspectiva, la discusión no se sitúa en decidir si un texto escrito es o no multisemiótico, sino en saber cuántos sistemas semióticos se consideran en lo “multi” y qué es un “sistema”. Estas constituyen preguntas primordiales al aproximarnos a este ámbito relativamente nuevo en la indagación científica del texto escrito.

Más recientemente, desde el ámbito lingüístico, este tema es traído a escena por los trabajos de Kress y Van Leeuwen (1992, 1996, 2001), aunque con un término que ha adquirido gran difusión: **multimodalidad**. También el estudio de los géneros discursivos al indagar en la naturaleza de los géneros y los textos de las disciplinas científicas ha puesto de manifiesto la necesidad de estudiar la multisemiosis como rasgo inherente a los textos, superando así la mirada más exclusivamente lingüística desde la tipologías textuales. Junto a ello, las indagaciones de los textos y los géneros emergentes de la web, del mismo modo, han llamado la atención sobre los artefactos multisemióticos (Jewitt, 2009).

Así, Lemke (1998), muy acertadamente, afirma que no se puede construir significado por medio de un solo y exclusivo sistema semiótico. Filogenéticamente, la comunicación humana ha desarrollado progresivamente el empleo conjunto de múltiples sistemas semióticos para crear cada vez más complejos y completos discursos que permitan la construcción de conocimiento avanzado, así como su transmisión y acceso a través de diversos géneros. De hecho, el mismo Lemke (2005) sostiene que los géneros, en ciertos contextos muy contemporáneos, pueden llegar a diluirse y se hibridizan de modo muy flexible, llegando a constituirse en construcciones que transitan ya no institucionalmente, sino temática o actitudinalmente.

Por su parte, Bateman (2008) es enfático en señalar que sus aportes ofrecen una nueva dirección a los estudios en multimodalidad y géneros del discurso, por lo que afirma que con ello se abre un mundo que el autor denomina era “post-Kress & van Leeuwen” (Bateman, 2008: XIX). Aunque Martinec y van Leeuwen (2009) no escatiman creatividad al abordar el estudio de las denominadas “traducciones” entre un texto más predominantemente verbal a otro en formato de página web de tipo muy dinámico. Hecho que pone en duda la afirmación de Bateman (2008).

Como bien señala O’Halloran (2005), muy posiblemente debido a lo relativamente nuevo del ámbito, aún no se cuenta con una terminología ampliamente consensuada en la comunidad disciplinar, hecho que ha llevado a ciertas confusiones y falta de precisiones. Tal vez lo más serio resulta la confusión entre los términos *modo* y *modalidad*, ambos con el sentido de *sistema semiótico*. Así, desafortuna-

damente, algunos autores han derivado del término *multimodalidad* la raíz *modo* (Kress & Van Leeuwen, 1996, 2001; Stöckl, 2004; Royce, 2007; Bateman, 2009; Kress, 2009; Stein, 2009) y hablan, por ejemplo, del “modo gráfico o visual” y del “modo verbal” del lenguaje o del texto. Incluso, Kress (2009) y Stein (2009) sostienen que, por ejemplo, el modo gráfico es otro tipo de modo tal como lo es el oral o el escrito. Desde la correcta derivación y técnicamente hablando, correspondería decir “modalidad gráfica o visual” y “modalidad verbal” o “sistema semiótico gráfico o visual” y “sistema semiótico verbal”. En cambio, entre otros, desde la Lingüística Sistémico-Funcional, el término *modo* ha venido empleándose preferentemente para referirse al canal de comunicación (e.g., escrito u oral) y, como se comprende, no para referirse al tipo de sistema semiótico (e.g., verbal, gráfico, matemático o tipográfico). En nuestra opinión, esta distinción es fundamental pues no se está enfocando el mismo objeto ni el mismo nivel de análisis. Por ello, con el fin de mantener una separación no solo terminológica sino primordialmente de tipo conceptual, es más acertado hablar de *multisemiosis* y, consecuentemente, de *texto multisemiótico*. En todo caso, se debe reconocer que el término *multimodal* ha alcanzado una amplia difusión, ha ganado un espacio importante y ha creado su propio ámbito de estudio. La polisemia que radica en este término no es ajena tampoco al término *modalidad*, el cual –tal vez desafortunadamente– en la literatura de la especialidad también tiene más de un sentido técnico.

Ahora bien, con el fin de contar con una terminología clara y precisa, en esta investigación utilizamos el término *multisemiosis* o *multimodalidad* para referirnos a la organización de los diversos sistemas semióticos que dan forma a un texto escrito. Por su parte, reservamos el término *modo* para la distinción –entre otros– de oralidad y escritura.

1.2.1. Los textos escritos (multisemióticos): ¿Qué es un sistema o modalidad y cuántos existen?

En el marco de esta investigación con foco en el texto académico disciplinar, sostenemos que un “sistema semiótico” o “modalidad” está constituido por un conjunto o repertorio de signos de una misma naturaleza (e.g., verbal, gráfica, matemática, etc.) que se articulan interrelacionadamente a partir de principios de organización funcional, semántica o morfológica, propios a cada sistema; en otras palabras, un sistema semiótico cuenta, a la vez, con un grupo de unidades constitutivas y un tipo de sintaxis que regula su disposición. Estos sistemas también se caracterizan por una ordenación específica de las unidades, ordenación que responde a una forma de conceptualizar y significar, haciendo así uso singular del espacio dentro del texto. De modo concreto, en el caso del texto escrito estático, estos sistemas se construyen para servir propósitos comunicativos según las necesidades de los escritores, dentro de un

contexto social y cultural determinado. En este sentido, estos sistemas o modalidades pueden realizarse a través de diferentes medios de representación y pueden presentar relaciones intersemióticas de diverso tipo (Stöckl, 2004; Kong, 2006; Royce, 2007; Matthiessen, 2007; Jewitt, 2009; Bednarek & Martin, 2010).

Royce (2007) reconoce que se han producido escasos trabajos que vinculen y expliciten la naturaleza de las relaciones semánticas intersemióticas entre los sistemas verbal y visual. En particular, Royce (2007) pone énfasis en buscar la explicación de los rasgos que hacen que un texto sea verbal-visualmente coherente. En este punto, dado que Royce (2007) al igual que otros investigadores (e.g., Bateman, 2008; Martin & Rose, 2008; Bednarek, & Martin, 2010) intentan proponer marcos de estudio desde el análisis del discurso, en nuestra opinión, llama poderosamente la atención cómo se cruza una línea entre lo que corresponde puramente a la descripción de los rasgos de un texto multisemiótico y lo que atañe directamente al procesamiento psicosociolingüístico de un texto multisemiótico entre un hablante/escritor y un oyente/lector. En otras palabras, lo que resulta curioso es que desde un análisis exclusivamente descriptivo, realizado solo a partir de algunos textos, se llega a establecer conexiones entre una modalidad y otra sin que se indague empíricamente en el procesamiento efectivo por parte de un lector. Esta cuestión resulta altamente relevante para el estudio de la multisemiosis y de las relaciones intersemióticas, pues –en nuestra opinión– falta un estudio de tipo psicolingüístico que pueda aportar efectivamente a comprender cómo un lector establece las relaciones entre las modalidades y de qué tipo son estas vinculaciones. Sin lugar a dudas, este nicho emergente y escasamente explorado se abre como un espacio fecundo de investigación, el que en nuestro proyecto mayor de investigación buscamos abordar desde un enfoque “basado en el corpus”.

Ahora bien, desde el objeto específico que nos preocupa, es decir, el *texto escrito estático* a través de un grupo de disciplinas científicas, partimos de la base de que existen cuatro sistemas semióticos o modalidades que se complementan e interactúan sinérgicamente para construir significados en un texto. Ellos son el **sistema verbal**, el **sistema gráfico**, el **sistema matemático**, y el **sistema tipográfico** (Lemke, 1998, 2003; O’Halloran, 2004, 2005, 2008; Van Leeuwen, 2005, 2006; Bednarek & Martin, 2010). Es fundamental establecer que, aunque se propone esta distinción cuatripartita, estos cuatro sistemas se articulan a través de los géneros y las disciplinas, dando origen a múltiples artefactos en una indeterminada cadena de posibilidades emergentes; todo ello con el fin de construir complejos significados.

La Figura 1 muestra esta concepción del texto escrito y da cuenta de sus cuatro sistemas constitutivos.

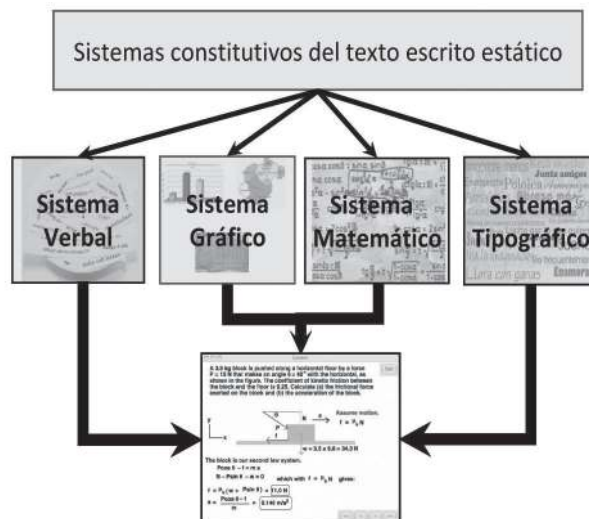


Figura 1. Cuatro sistemas semióticos en el texto escrito estático.

El sistema verbal está conformado por palabras, frases y oraciones que constituyen la expresión de significados basados en lo exclusivamente lingüístico. Esto quiere decir que la materia de expresión es la lexicogramática. Así, existe un léxico y una sintaxis que en cada lengua se articula de modo singular.

Por su parte, el sistema gráfico se constituye a partir de trazos que dan forma a una representación pictórica de información de diversa índole, la cual permite la presentación de datos en determinados formatos. Esta representación posibilita comunicar mediante medios gráficos tanto ideas, hechos y valores procesados y sintetizados en términos de formato visual, así como también factores sociales, culturales, económicos, estéticos y tecnológicos. Hoy en día, gran parte del diseño gráfico es asistido por herramientas digitales a través de computadores, hecho que ha venido a enriquecer las posibilidades y medios disponibles. Las fotografías, gráficos, diagramas, tablas y bocetos son recursos típicos del lenguaje gráfico, en algunos de los cuales paralelamente se suele combinar más de un modo. Conjuntamente, el uso de los espacios en blanco también se constituye en un recurso importante en el modo gráfico, pues el adecuado equilibrio de su empleo es parte de la composición visual.

El sistema matemático está conformado por grupos de grafías, signos o representaciones de diversa índole o también denominados alfabetos, los cuales permiten que el significado sea codificado simbólicamente de modo sintético. Los elementos constitutivos de este sistema interactúan a partir de una o varios tipos de sintaxis. De modo más preciso, este sistema –entre otros– está compuesto por números (romanos o árabes), letras (griegas o latinas), operadores y signos de pun-

tuación. A través del desarrollo filogenético de la especie humana, este sistema no ha alcanzado una evolución tal que le permita autonomía total, por ello requiere de los sistemas verbal y tipográfico –entre otros– en sus co-textos con el fin de construir una intersemiosis que conlleve los significados completos que se busca comunicar.

Por último, el sistema tipográfico se constituye por la forma y color de las letras, el cual incluye –entre otros– negrita, cursiva, alta (mayúscula), baja (minúscula), color, dimensionalidad (uni, bi o tridimensional). Los significados tipográficos como parte del significado total del texto escrito han sido posiblemente los más olvidados y frente a los cuales ha existido escasa conciencia por parte de los lingüistas (Van Leeuwen, 2005, 2006). La modalidad tipográfica no constituye un recurso de vestidura, ropaje o simplemente decorado, pues no se restringe a una cuestión de estética visual; así, a través de sus diversos recursos, este sistema aporta un potencial de significado que –en conjunto con otros sistemas– puede llegar a transmitir de esta forma el significado global del texto. Ciertamente en sus orígenes este sistema tipográfico puede haber emergido para cumplir un propósito ornamental o decorativo, pero –en su evolución– los seres humanos lo han llevado a desempeñar un rol mucho más dinámico. De hecho, con el devenir de los procesadores de textos computacionales, los escritores han tomado mucha mayor conciencia de la riqueza que reside en su aporte y así ha pasado a desempeñar un rol mucho más activo.

1.2.2. El concepto de texto (multisemiótico)

Resulta importante precisar que en nuestra concepción del texto escrito los cuatro sistemas semióticos descritos son constitutivos en su naturaleza, aunque no necesariamente todos siempre deban estar presentes. Es decir, en la construcción de los significados que se elaboran en el texto, los cuatro se suelen articular en conjunto para elaborar y comunicar el significado textual. Con lo ilustrado en la Figura 1 no se intenta proponer que estos sistemas son pre-existentes al texto en sí mismo, sino que los seres humanos los hemos construido y desarrollado a través del tiempo para dar cuenta de los significados que deseamos comunicar por medio de los textos. Esto es, ellos son consustanciales a los textos escritos y a los significados que los seres humanos hemos sentido necesidad de expresar en contextos y momentos específicos, es decir, cumplen un propósito específico. Desde esta perspectiva, sostenemos que no es exacto decir que “existe el texto y las figuras” o que “existe el texto y los artefactos multisemióticos”. Muy por el contrario, se debe enfatizar firmemente que existe el texto y en él se suelen conjugar sinérgicamente el sistema verbal, el gráfico, el matemático y el tipográfico; consecuentemente, según sea el caso de los significados comunicados, existirá mayor predominancia o aparición de unos y otros. Así, es posible que existan textos en que no todos estos cuatro

sistemas estén necesariamente presentes, justamente las disciplinas y los géneros dan cuenta de ello: mayor presencia y predominancia de ciertos sistemas por sobre otros. Ahora bien, resulta muy significativo dejar constancia que estas precisiones tienen que ver con la concepción de texto que propugnamos y su naturaleza eminentemente multisemiótica. Es posible que, desde la lingüística, la hegemonía del estudio del sistema verbal por sobre otros haya llevado a estas expresiones –en mi opinión– erróneas o a estas aparentes confusiones, por lo cual es primordial explicitar la concepción que fundamenta nuestra investigación y deslindarla de estas apreciaciones o concepciones menos precisas.

En este sentido, también es prudente resaltar que cada uno de estos sistemas semióticos presenta relaciones intrasemióticas a la vez que intersemióticas, tal como se aprecia en la Figura 2.

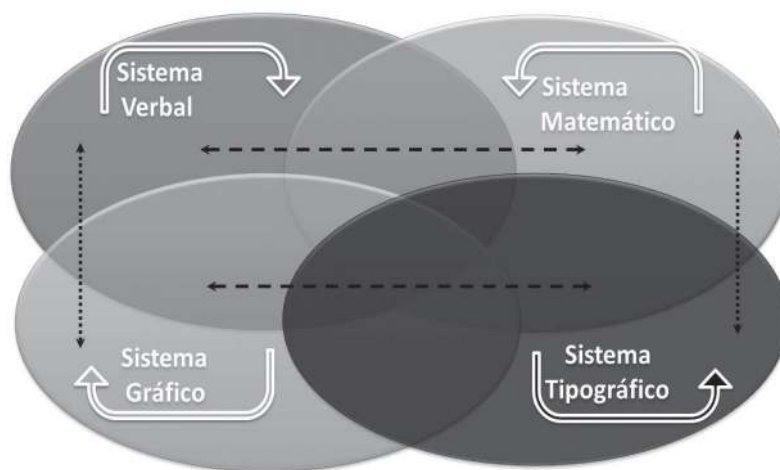
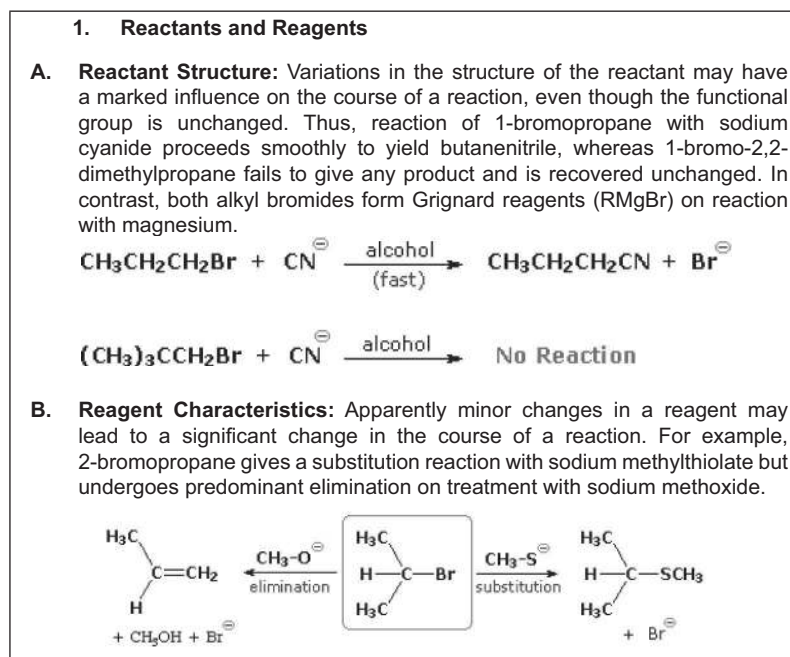


Figura 2. Los sistemas semióticos y sus relaciones.

Lo que se intenta capturar en esta figura es, por un lado, la estrecha vinculación de unos sistemas con otros y, por otro, lo complejo de intentar –en algunos casos– una separación entre un sistema semiótico y sus recursos y otro u otros sistemas. Por ejemplo, cuando se trata de definir el sistema matemático, aunque pudiera pensarse que éste existe como un sistema autónomo y que resultaría muy sencillo deslindarlo del verbal, del gráfico y del tipográfico, la realidad que impone el análisis de un determinado texto revela una complejidad intrínseca. Ello no constituye en sí mismo un problema, en cambio lo que hace es poner de relieve la naturaleza misma de la multimodalidad textual y la sinergia de los sistemas interrelacionados. De hecho, el sistema matemático en parte se apoya en el co-texto verbal y hace uso explícito de los sistemas gráfico y tipográfico. En efecto, a modo de ejemplo, véase

el siguiente extracto tomado de un texto de química del corpus en estudio:



Como se comprende, a partir del ejemplo anterior, las cuatro modalidades señaladas no pueden ser separadas radicalmente ni concebidas como externas al texto, salvo por razones metodológicas, operacionales o de foco de investigación. Todas ellas, en este caso, se despliegan con diversas relaciones intersemióticas y en conjunto dan forma a los significados de este pasaje de un texto del área de química. Y también se aprecia con mucha claridad lo que se comentaba a partir de la Figura 2, en donde se despliegan las relaciones intrasemióticas hacia el interior de un sistema así como las diversas posibles vinculaciones intersemióticas entre dos, tres o los cuatro sistemas.

1.2.2. El desafío de la automatización

Desde las investigaciones que siguen principios de la lingüística de corpus (Parodi, 2010a), el reconocimiento automático de los artefactos multisemióticos a través de programas computacionales tanto para textos estáticos como dinámicos surge como un reto inminente. Si, como se ha señalado insistentemente, los estudios del lenguaje han propendido a una hegemonía de lo lexicogramatical, consecuentemente la mayoría de las herramientas computacionales disponibles de análisis automático de textos se ha limitado a la información lingüística. Esto significa

que las tablas, diagramas, gráficos y fórmulas, sólo por mencionar algunos artefactos multisemióticos, así como los diseños y disposición en la página, no se han considerado de modo sistemático en el análisis de corpus, a pesar de que la mayoría de los géneros en casi todas las disciplinas científicas están involucrados con textos multisemióticos (Halliday & Martin, 1993; Lemke, 1998; Martin & Veel, 1998; Matthiessen, 2007; Wignell, 2007; Martin & Rose, 2008; Parodi, 2010b y 2010c).

Aunque con ciertas leves diferencias, Bateman, Delin y Henschel (2004) y Baldry (2007) abogan en la misma línea que Parodi (2010c), esto es, en dirección a realizar indagaciones empíricas en la perspectiva de lo que Parodi (2010c) denomina una “lingüística de corpus multisemiótica” (LCM). Si bien esta denominación podría resultar explícitamente inoportuna, a la luz de lo discutido en el apartado precedente, pues ello podría estar de más o llegar a ser una sobreexplicación de lo obvio. Así, si un corpus está compuesto de textos y –como se ha defendido aquí– los textos son intrínsecamente de naturaleza multisemiótica, el término multisemiótica en LCM podría estar de sobra; aunque también se podría argumentar que su inclusión es aclaradora.

No obstante lo complejo de la tarea del reconocimiento automático, algunos avances importantes en la investigación de textos multisemióticos se destacan (Delin, Bateman, & Allen, 2002/3; Kong, 2006; O’Halloran, 2005, 2008). Por ejemplo, en el Proyecto “Género y multimodalidad: Un modelo computacional de género en el documento de diseño” (GEM) (Delin, Bateman, & Allen, 2002/3) se llevó a cabo una visión multisemiótica de género con el objetivo de producir un esquema notacional que permitiera la descripción automática de varias capas de documentos ilustrados con diseño complejo. Más precisamente, se intentaba establecer empíricamente el grado de relación sistemática y regular entre algunos géneros (por ejemplo: manuales de instrucciones, periódicos, libros ilustrados, páginas web) y sus realizaciones en los textos complejos que incluyen, en conjunto, formatos verbales y visuales tales como diagramas, imágenes y gráficos. En el actual proyecto GEM, Bateman (2008) busca ambiciosamente integrar tres enfoques analíticos: partiendo con un análisis visual, el cual luego se integra a uno de tipo retórico (Teoría Retórica Estructural) y se completa con la teoría del género con base en la Lingüística Sistemico-Funcional. El desafío reside, y esto es a lo que se apuesta más bien para el futuro, a que este análisis se lleve a cabo automáticamente por un computador.

Paralelamente, en el proyecto “Laboratorio de Análisis Multimodal” en la Universidad de Singapur (O’Halloran, 2004, 2005, 2008; O’Halloran, Tan, Smith & Podlasov, 2010) existe un equipo de investigadores de ciencias sociales y ciencias de la computación que trabajan colaborativamente para desarrollar un prototipo de software para modelar, analizar, almacenar y recuperar el sentido de los textos a partir de imágenes, textos interactivos de vídeo y los sitios digitales, los que se

construyen a través de la utilización de múltiples recursos semióticos (e.g., lenguaje, imágenes visuales, gestos, movimientos, música, sonido, objetos tridimensionales, entre otros). Este equipo se encuentra interdisciplinariamente explorando la compleja dinámica de las prácticas integrales de construcción de significados (<http://multimodal-analysis-lab.org/>).

Por su parte, los esfuerzos desplegados por Baldry y Thibault (2001; Baldry, 2007) en el desarrollo de un sistema computacional en línea, basado en lenguaje XML (*Extensible Markup Language*), para un buscador de concordancias multisemióticas (denominado *Multimodal Corpus Authoring*, MCA) debe considerarse como un hito altamente relevante. Este recurso apoya el trabajo de investigadores con textos dinámicos de tipo películas, páginas web y otros similares.

En este contexto, enfrentar el análisis multisemiótica automático de un corpus requiere el desarrollo de sofisticadas herramientas computacionales y aun se perfila como un desafío a futuro (Sköckl, 2004; Bateman, 2008; Parodi, 2010c). Un notable adelanto lo constituyen los lenguajes tipo SGML (*Standard Generalized Markup Language*) y XML (Bryan, 1988; CES, 2000). Ellos han mostrado —en principio— ser recursos extremadamente valiosos para identificar automáticamente algunos de los aspectos semióticos en el texto. Estas herramientas ofrecen normas preliminares y marcos de anotación de corpus, pero en la actualidad no garantizan un proceso totalmente automático de artefactos multisemióticos con alta precisión y sólida consistencia y corrección. Es más, la identificación y anotación digital automática de textos en formato electrónico carece de una teoría integral del lenguaje (multisemiótico) suficientemente robusta en el marco del denominado “giro visual” (Parodi, 2010c). No obstante nuestra visión crítica, un grupo importante de investigadores en esta área ha visto en la Lingüística Sistémico-Funcional de la Escuela de Sydney un referente importante para estas indagaciones. El problema reside en que su marcado anticognitivismo hace compleja la comprensión profunda de un objeto cuyo procesamiento psicolingüístico junto a su necesaria representación mental constituyen núcleos fundamentales en esta indagación. En nuestra opinión, el análisis descriptivo que se pueda hacer de los textos siempre deberá ser corroborado por la investigación psicolingüística de dicho procesamiento, lo que en definitiva aportará los datos concluyentes.

2. LA INVESTIGACIÓN: CORPUS Y METODOLOGÍA

2.1. Recolección del Corpus PUCV-2010

El Corpus PUCV-2010 se compone de los materiales de lectura de estudiantes de doctorado en las áreas de Ciencias Sociales y Humanidades y Ciencias Básicas. El corpus fue recogido en doce programas de doctorado (dos por especialidad) en seis

universidades chilenas y comprende todos los documentos que se registran como material de lectura en las asignaturas obligatorias de sus respectivos programas de estudio, con la excepción de los incluidos en la investigación doctoral final. En la Tabla I se da cuenta de las seis disciplinas científicas seleccionadas para constituir el corpus en estudio.

Tabla I. Áreas y disciplinas seleccionadas.

Área Científica	Disciplina
Ciencias Sociales y Humanidades	Historia
	Lingüística
	Literatura
Ciencias Básicas	Biología
	Física
	Química

Estas disciplinas fueron seleccionadas con el objetivo de obtener una panorámica amplia de posibles variaciones de los documentos escritos provenientes de los respectivos programas de educación doctoral. Se espera así llegar a contar con un corpus diversificado que sustente estudios variacionistas en diversas líneas.

El procedimiento inicial de recolección y digitalización de los textos del corpus en esta fase de la investigación ha sido relativamente similar al que hemos desarrollado en otras investigaciones precedentes (Parodi, 2010b) y se resume en la tabla siguiente.

Tabla II. Recolección y procesamiento del corpus PUCV-2010.

Paso 1: Construcción de una base de datos con la información acerca de las universidades que imparten los seis programas de doctorado en estudio en Chile
Paso 2: Contacto con los directores de los programas seleccionados con el fin de solicitar los programas de estudio de todas las asignaturas del plan académico de cada doctorado
Paso 3: Construcción de una base de datos con la información completa de los currículos de los programas de doctorado en las seis especialidades y sus respectivos programas de cada asignatura

Continuación Tabla II.

Paso 4: Construcción de una base de datos con todas las referencias bibliográficas obligatorias y de consultas, incluidas en dichos programas de estudio
Paso 5: Búsqueda en Internet con el fin de encontrar aquellos títulos disponibles en formato digital, minimizando así el tiempo de digitalización
Paso 6: Recolección de los textos de las bibliotecas correspondientes y de las oficinas de los profesores
Paso 7: Entrenamiento de un equipo de asistentes para escanear y compilar todos los textos
Paso 8: Procesamiento y conversión de todos los textos del corpus en formato PDF editable y en formato plano (*.txt)
Paso 9: Carga de cada texto a través del programa El Grial (www.elgrial.cl) en la plataforma en línea junto a su correspondiente procesamiento de etiquetado morfosintáctico

2.2. Constitución del Corpus PUCV-2010 y la muestra en estudio

Basados en los pasos 1 al 7 de la tabla anterior, hemos llegado a construir la siguiente cuantificación de los textos del corpus. Esto se recoge en la Tabla III.

Tabla III. Constitución numérica del Corpus PUCV-2010.

Área y Disciplina		Nº de Textos por Disciplina	Totales por Área
Ciencias Sociales y Humanas	Lingüística	1.075	2.616 (82,8%)
	Literatura	946	
	Historia	595	
Ciencias Básicas	Biotecnología	227	544 (17,2%)
	Física	181	
	Química	136	
Total		3.160 (100%)	

Debido al tamaño del corpus y el tiempo requerido para contar con una total disponibilidad del material para el estudio de los artefactos multisemióticos, decidimos realizar un estudio con una muestra aleatoria y proporcional del corpus en construcción al 33% del total. En la Tabla IV se muestra la constitución numérica

de esta muestra a partir del corpus de referencia.

Tabla IV. Constitución de la muestra en estudio.

Disciplinas	Total de textos del Corpus	Muestra aleatoria y proporcional (33%)
Biotecnología	227	75
Física	181	60
Química	136	45
Lingüística	1.075	355
Literatura	946	312
Historia	595	196
Total	3.160 (100%)	1.043 (33%)

Para la selección de los textos de la muestra se empleó un programa computacional *ad hoc* provisto de una fórmula matemática, el cual permitió la selección aleatoria del número requerido de textos a partir del listado total de cada subcorpus por disciplina científica. De este modo, llegamos a la muestra de 1.043 textos, divididos por disciplina según se constata en la Tabla IV. En este punto, vale la pena recordar, tal como se indicó en la Introducción, que en este estudio se atiende al discurso académico y su caracterización desde los recursos multisemióticos empleados a través de las seis disciplinas en estudio. Esto quiere decir que en esta fase de la investigación no se atiende a la variación de artefactos multisemióticos por géneros discursivos; ello será materia de otro estudio posterior.

3. HALLAZGOS Y RESULTADOS

3.1. Identificación y definición de los artefactos multisemióticos en el Corpus PUCV-2010

Con el fin de pesquisar los artefactos presentes en la muestra de textos y llevar un riguroso conteo de la ocurrencia de cada artefacto en los textos de las seis disciplinas en estudio, se entrenó a un equipo de asistentes que manejaran los principios teóricos y de los criterios operacionalizados para su identificación. Luego de una fase de entrenamiento y de calibración del empleo de los criterios en aplicación, se comprobó que los asistentes lograron un alto desempeño en la identificación y conteo de modo equilibrado y sistemático. Esto se calculó en base a un trabajo con un grupo de textos de entrenamiento, en los cuales se alcanzó un grado de

validación o fiabilidad entre pares expertos por sobre el 98% de acuerdo.

Ahora bien, cabe precisar que nuestro interés de estudio en esta primera parte de la investigación no incorpora la indagación acerca de las interacciones entre estos cuatro sistemas semióticos ni de la forma en que ellos se complementan o “traducen” de una modalidad a otra (como sí le interesa –entre otros– a Royce, 2007; Bateman, 2008; Martinec & Van Leeuwen, 2009; y O’Halloram, 2010).

Basados en los cuatro sistemas semióticos ya identificados (Verbal, Gráfico, Matemático y Tipográfico) fue posible identificar nueve artefactos multisemióticos presentes en la muestra de 1.043 textos del Corpus PUCV-2010. Como se puede apreciar, decidimos utilizar nombres relativamente comunes que corresponden a lo que tradicionalmente hemos empleado para referirnos a estos artefactos. Tal vez, el artefacto más novedoso sea el que hemos denominado **Red Composicional**, que más adelante definimos. Estos nueve artefactos se listan en orden alfabético en la Tabla V.

Tabla V. Identificación de artefactos multisemióticos.

Artefactos Multisemióticos Corpus PUCV-2010
Esquema
Figura Geométrica
Fórmula
Gráfico Estadístico
Ícono
Ilustración
Mapa
Red Composicional
Tabla

Una primera cuestión muy relevante es apuntar que, dado el muestreo al 33%, es muy posible que en un posterior análisis al resto de los textos del corpus emerjan nuevos artefactos no capturados en este estudio preliminar. Por ello, los resultados cuantitativos entregados más adelante también deben tomarse en su carácter exploratorio y aproximativo dentro de la descripción del Corpus PUCV-2010.

Ahora bien, cabe destacar que cada uno de estos artefactos presenta límites porosos o difusos en su constitución, es decir, bajo cada nombre se ha identificado y agrupado a un conjunto o más bien a un rango o continuum de recursos. Ello, dado que partimos del supuesto que estas categorías no son dicotómicas ni discretas en su totalidad. Las siguientes figuras 3 y 4 ejemplifican este principio a partir del artefacto **Esquema** y del artefacto **Gráfico Estadístico**.

Esquemas

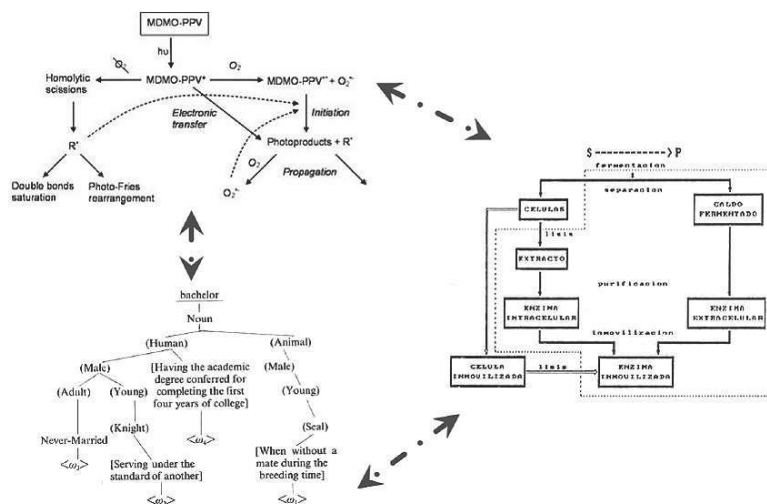


Figura 3. Variación del artefacto Esquema.

Como se aprecia en esta figura, se muestran 3 tipos de esquemas. Se observa que los tres caben dentro de la categoría Esquema, pero la variabilidad resalta sus posibilidades de realización en diversos casos concretos. Algunos podrían considerarse como casos más prototípicos y otros más en el borde de la clasificación, pero todos ellos comparten su pertenencia a la categoría nominada como **Esquema**.

Por su parte, situación similar se detecta en el caso del artefacto Gráfico Estadístico. Veamos los ejemplos de la Figura 4.

Gráficos Estadísticos

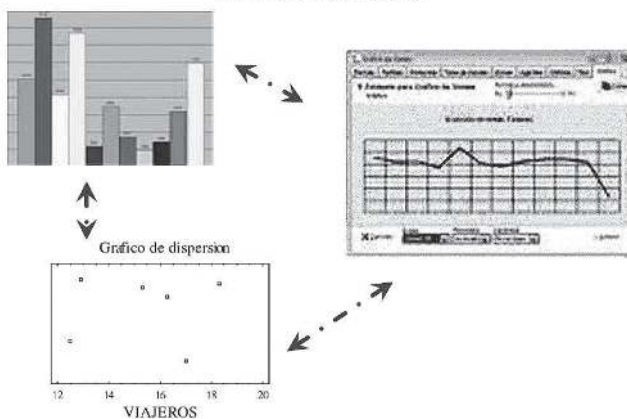


Figura 4. Variación del artefacto Gráfico Estadístico.

Con el fin de precisar de modo más certero cada uno de estos artefactos, procedemos a la construcción de una definición operativa que dé cuenta de ellos de manera más o menos sistemática. Para ello se seleccionaron tres criterios: modalidad, función y composición. Ellos brindan un modelo para elaborar cada una de las nueve definiciones. En la Tabla VI se listan los criterios y la pregunta orientadora para comprender el criterio, los cuales se plasman en el mismo orden en cada definición.

Tabla VI. Criterios empleados en la construcción de las definiciones.

Criterio 1	Criterio 2	Criterio 3
Modalidad (¿Qué sistemas participan en el artefacto?)	Función (¿Para qué se emplea el artefacto?)	Composición (¿De qué se constituye el artefacto?)

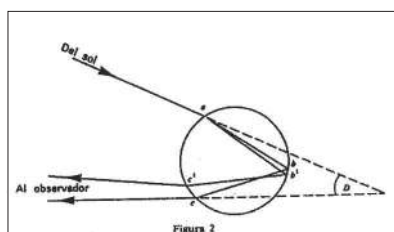
En lo que sigue se entrega una definición y se ejemplifica cada artefacto multisemiótico con muestras tomadas de los textos del Corpus PUCV-2010. Ellos se resumen en la Tabla VII.

Tabla VII. Definición y ejemplificación de los 9 artefactos multisemióticos.

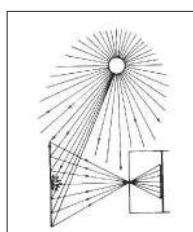
1. Esquema: artefacto que suele combinar la modalidad verbal, la gráfica y la tipográfica. Esta representación de significados da cuenta sintéticamente de un constructo teórico o empírico y, generalmente, muestra resumidamente la organización y vinculación jerárquica entre partes o componentes de un objeto o concepto determinado. Los esquemas pueden construirse –entre otros– a partir del uso de cuadros, operadores funcionales o relacionales (e.g., flechas, paréntesis cuadrados, de llaves, etc.), viñetas, numeraciones, tamaño y tipo de letras (e.g., mayúsculas y minúsculas; negritas y cursivas) y colores.

BIO16 **LIN1.268**

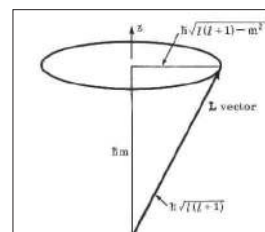
2. Figura Geométrica: artefacto que utiliza de manera preferente el sistema gráfico, aunque en menor medida también se suele combinar con el sistema verbal, el tipográfico y el matemático. Es, mayoritariamente, la representación de un espacio encerrado entre líneas que busca dar cuenta de un objeto de la realidad o de un constructo teórico. Sus componentes principales son el plano, el punto, la línea –recta, curva, quebrada–, la superficie, y el segmento. Mediante transformaciones y desplazamientos de sus componentes generan diversas líneas, superficies y volúmenes; así, ellas pueden tener formas uni o multidimensionales (bidimensionales, tridimensionales, etc.). Las figuras geométricas más comunes son –entre otros– los cuadrados, rectángulos, círculos, pirámides y polígonos.



LIT3



LIT183



FIS109

3. Fórmula: artefacto que se construye preferentemente a partir de tres modalidades: la matemática, la verbal y la tipográfica. La fórmula permite expresar de forma breve información que relaciona cantidades, constantes o variables. Ella incluye una cadena de caracteres, organizados a partir de una sintaxis y con base en uno o varios alfabetos. Esta representación puede incluir números romanos o árabes, letras griegas o latinas (en mayúsculas o minúsculas), operadores (funciones) y signos de puntuación. La tipográfica se revela –entre otros– por medio de las negritas y las mayúsculas.

$$\int \mathbf{F} \cdot d\mathbf{s} = m \int \frac{d\mathbf{v}}{dt} \cdot \mathbf{v} dt = \frac{m}{2} \int \frac{d}{dt} (v^2) dt,$$

FIS127

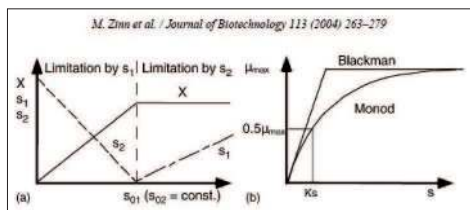
$$|\vec{F}| = \frac{|q_1 q_2|}{r^2}.$$

FIS108

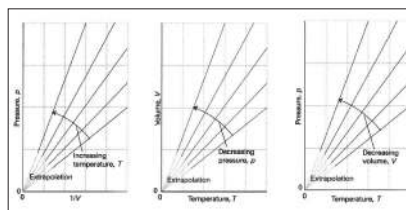
$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \quad \text{y} \quad \sum_{n=0}^{\infty} \frac{1}{n!}$$

LIN1.035

4. **Gráfico Estadístico:** artefacto que combina preferentemente las cuatro modalidades: verbal, gráfica, matemática y tipográfica. En él se representa pictóricamente un resumen visual de información estadística. Los gráficos estadísticos pueden ser –entre otros– de líneas, de barras, de tortas, o de tipo histogramas.



BIO25

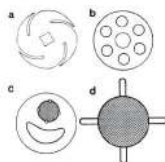


QUI49

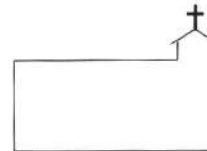
5. **Ícono:** artefacto que preferentemente se construye solo a partir de la modalidad gráfica. Representa o sustituye mayoritariamente un objeto real o ideal y, cuando procede, suele tener alta semejanza con un objeto en el mundo real. También representa de forma rápida y fácil un concepto ideal o empírico o asociación, como sucede, por ejemplo, en la informática (e.g., un programa, una función, un documento o un archivo o carpeta). Tradicionalmente eran solo figuras o dibujos planos (unidimensionales), pero en la actualidad pueden ser multidimensionales.



QUI49



BIO53



LIN1.180

6. Ilustración: artefacto que suele emplear preferentemente la modalidad gráfica, aunque en menor medida también se combina con la verbal y tipográfica. Principalmente, representa –mediante técnicas de fotografía, dibujo, estampado o grabado– objetos reales o ideales. Normalmente se considera que adorna o complementa, ya sea en blanco y negro o en colores, otra parte del texto que suele ser verbal. Se compone de trazos libres o también se dice que suelen capturar de modo certero o directo una porción de la realidad, permitiendo registrarla y posteriormente reproducirla; así, la ilustración registra con cierto grado de precisión circunstancias, fenómenos, estados o acontecimientos tradicionalmente en forma plana, aunque hoy en día también puede efectuarse de modo multidimensional.



HIS301



Fachada del Club Naval, Valparaíso

HIS7

7. Mapa: artefacto que preferentemente se construye a través de una sola modalidad: la gráfica. Sin embargo, también puede combinar el modo verbal y el matemático. El mapa es una representación geográfica de una parte o de toda la superficie de la Tierra, en la que se pueden también detallar la localización, características de fenómenos físicos y sociales. Se suele distinguir mapas físicos, geográficos, mapamundi y de tipo estático y dinámico.



HIS268



HIS606

8. Red Composicional: artefacto que habitualmente combina las cuatro modalidades: verbal, gráfica, matemática y tipográfica. Representa sintéticamente información de diversa índole en una articulación original, dado que corresponde principalmente a composiciones complejas no-lineales emergentes a partir de los nuevos medios digitales (e.g.; captura de una página web). En su conformación composicional puede haber organizaciones jerárquicas particulares provenientes de los recursos prototípicos de los tres modos semióticos constitutivos.



QUI49



LIT609

9. Tabla: artefacto que combina preponderantemente los sistemas verbal, gráfico y tipográfico, aunque también suele integrar el cuarto sistema matemático. La tabla suele representar contenidos jerárquicos e información de tipo clasificatoria y definicional; también entrega resumidamente resultados y datos cuantitativos y cualitativos en una organización de tipo matricial.

Table 1.4 A paradigm for mood and tagging including function structure

	not tagged	tagged
declarative	He was <i>licking</i> the baby. Subject*Finite...	He was <i>licking</i> the baby, <i>wasn't he</i> ? Subject*FiniteFinite*Subject
interrogative	Was he <i>licking</i> the baby? Finite*Subject	Was he <i>licking</i> the baby, <i>was he</i> ? Finite*Subject . . .Finite Subject
imperative	<i>Tickle</i> the baby! -	<i>Tickle</i> the baby, <i>won't you</i> ? Finite*Subject

LIN1.268

TABLE 3. POSICIONAMIENTO EN FUNCIÓN DE LAS CUATRO DIMENSIONES: ADULTOS Y JOVENES

	Países Bajos 96	Francia 96	Alemania 96
I			
Adultos	79	72	71
Jóvenes	87	83	89
II			
Adultos	76	57	64
Jóvenes	64	65	73
III			
Adultos	67	86	60
Jóvenes	80	71	74
IV			
Adultos	70	45	52
Jóvenes	82	58	72

Fuente: Donald Lighter, «An End to European Immigration?», *APSL*, vol. LXI, núm. 3 (marzo 1987), pág. 92.

HIS193

A continuación pasamos a revisar los resultados de las cuantificaciones de estos recursos a partir de los 1.043 textos de la muestra en estudio.

3.2. Comparaciones generales y entre dominios disciplinares

Antes de dar inicio a los análisis descriptivos de orden cuantitativo, debemos rei-

terar que –como se ha declarado anteriormente– todos los cálculos que presentaremos corresponden a la muestra aleatoria y proporcional del 33% del Corpus PUCV-2010.

En primer lugar en el Gráfico 1 se puede observar la distribución porcentual de los nueve artefactos en su frecuencia de ocurrencia total a partir de los 1.043 textos analizados. Por supuesto, en este gráfico las seis disciplinas están tomadas en conjunto.

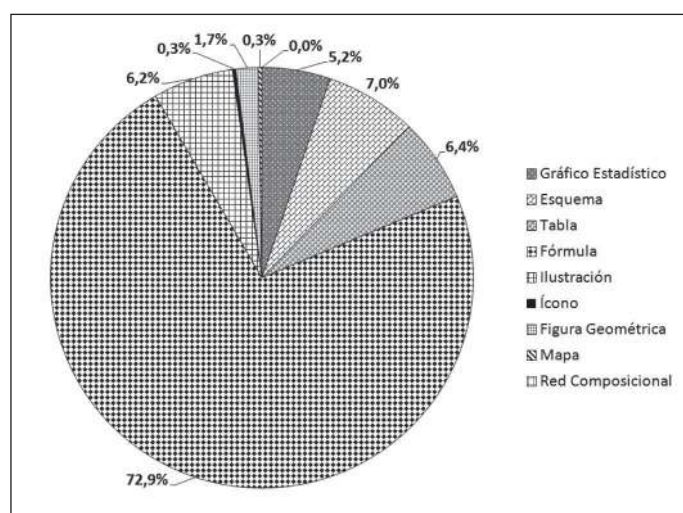


Gráfico 1. Distribución porcentual de artefactos en la muestra de textos.

Sin lugar a dudas, llama poderosamente la atención la alta ocurrencia porcentual del artefacto denominado **Fórmula**. Se hace evidente que en esta muestra, del total de artefactos multisemióticos identificados, el 72,9% corresponde a un artefacto que preferentemente se construye por medio de un sistema semiótico que es el lenguaje matemático, con apoyo secundario de los sistemas verbal, gráfico y tipográfico. Sin lugar a dudas, este hallazgo de tan elevada presencia de este artefacto aporta una certera caracterización preliminar del discurso académico especializado a nivel de estudios de doctorado. Muy posiblemente es factible hipotetizar que la ocurrencia de este artefacto no se realiza de manera homogénea a través de todos los textos de la muestra, en particular en las seis disciplinas en indagación. En los próximos análisis y gráficos aportaremos más información con el fin de dilucidar este asunto.

Muy distantes en porcentaje de ocurrencia respecto de la **Fórmula**, el **Esquema** (7%), la **Tabla** (6,4%), la **Ilustración** (6,2%) y el **Gráfico Estadístico** (5,2%) constituyen los cuatro artefactos que aparecen con las siguientes ocurrencias por-

centuales en la muestra total. Por una parte, la tremenda distancia estadística porcentual en la ocurrencia entre la **Fórmula** y los otros cuatro artefactos establece un hallazgo muy singular que pone de manifiesto la importancia de la **Fórmula** en el discurso académico a nivel de estudios doctorales. Se hace evidente que a través de ella y de las relaciones intersemióticas que se producen con los otros sistemas se construyen significados que ocupan una buena parte del material de lectura de los alumnos universitarios en educación superior. Por otro lado, también resulta interesante observar la singular homogeneidad en la ocurrencia porcentual de los otros cuatro artefactos ya mencionados. Por su parte, queda de manifiesto la escasa participación en el discurso académico, identificado en esta muestra, del **Mapa**, la **Figura Geométrica**, la **Red Composicional** y el **Ícono**. Por supuesto que su sola identificación brinda un hallazgo científico relevante; no obstante ello, también su ocurrencia porcentual tan baja dice mucho de su tipo de participación en la construcción y transmisión de significados científicos.

Luego de este ilustrativo primer acercamiento a estos datos, estudiamos la ocurrencia bruta (no porcentual) de los artefactos en la muestra, pero ahora separados entre Ciencias Sociales y Humanidades y Ciencias Básicas.

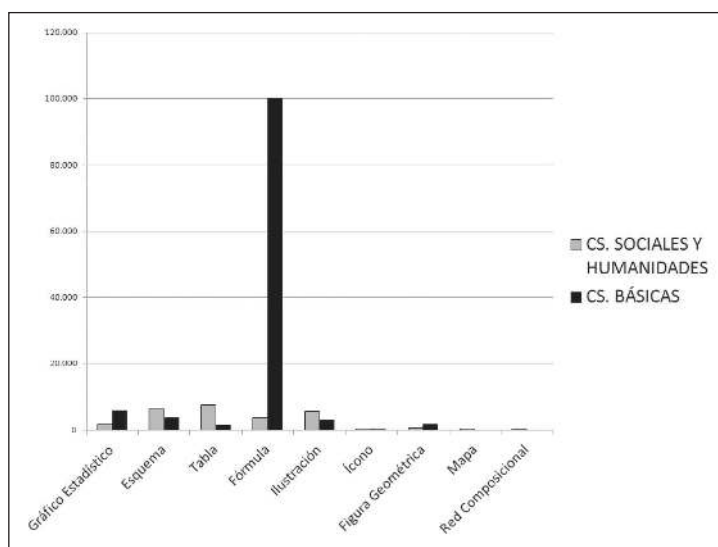


Gráfico 2. Ocurrencia bruta de artefactos en la muestra de textos.

Estos datos son ciertamente muy reveladores y posicionan claramente a la **Fórmula** dentro de los textos de la muestra y las disciplinas en un lugar muy singular. Tal como veníamos argumentando, este recurso semiótico se hace prototípico de la Ciencias Básicas, aunque también tiene presencia en los textos de las Ciencias Sociales y Humanidades. Por supuesto con una ocurrencia muy dispar. No obs-

tante ello, es sin lugar a dudas un hallazgo sustantivo por cuanto su presencia en el discurso disciplinar marca decididamente a algunas de las disciplinas en estudio.

Esto quiere decir que en los 545 textos que forman la muestra en Ciencias Básicas se identifica una ocurrencia de exactamente 100.124 fórmulas, mientras que en los 2.616 textos de la muestra de Ciencias Sociales y Humanidades se recogen 3.602 fórmulas; como se comprende, comparativamente el número de textos es considerablemente mayor en Ciencias Sociales y Humanidades, sin embargo, se hace evidente que este artefacto multisemiótico no es prototípico del discurso académico de estas áreas del saber. En cuanto a la relación número de artefactos y presencia por texto y página, en los próximos apartados retomaremos esto con cifras más precisas. Pero antes, demos una mirada a la distribución porcentual de los artefactos por dominio científico.

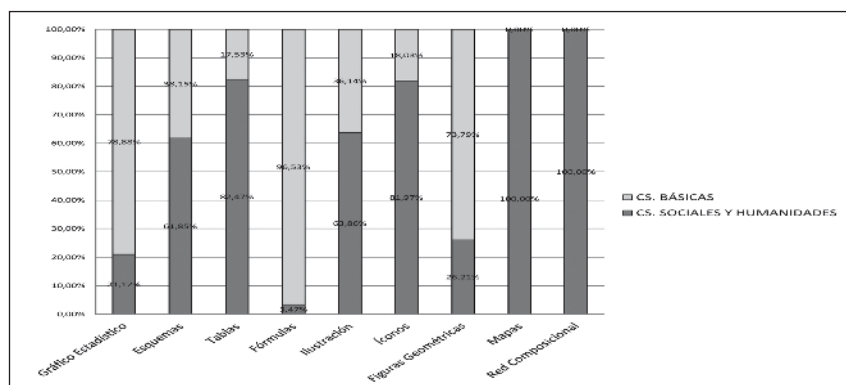


Gráfico 3. Distribución porcentual de artefactos en la muestra de textos.

La información aportada por el Gráfico 3 presenta algunos aspectos de modo más clarificador que el gráfico anterior. Por ejemplo, un asunto novedoso queda de manifiesto en cuanto a que el **Mapa** y la **Red Composicional** son artefactos exclusivos de las Ciencias Sociales y Humanidades, a diferencia de lo que ocurre con los otros siete artefactos, que –con ocurrencias mayores o menores– se registra su aparición en ambos dominios científicos. Junto a este hallazgo relativo a estos dos artefactos únicos en ocurrencia en un dominio científico, también debe destacarse el hecho de que la **Tabla** (82,47%), el **Ícono** (81,97%), la **Ilustración** (63,86%) y el **Esquema** (61,85%) se posicionan porcentualmente por sobre el 60% en comparación a las cifras de las Ciencias Básicas. Estos hallazgos, en una caracterización muy preliminar del discurso escrito de las Ciencias Sociales y Humanidades, pone de manifiesto importantes tendencias iniciales que dicen relación con el modo de construir significados disciplinares y el empleo de un amplio repertorio de recursos multisemióticos. Estos recursos y su ocurrencia relativamente homogénea

muestran –entre otros– los rasgos más prototípicos de los textos de las Ciencias Sociales y Humanidades, a los cuales los estudiantes de educación superior tendrían que acceder y a través de los cuales tendrían que aprender a comprender los contenidos del discurso académico, conocer sus géneros discursivos y a construir representaciones mentales dinámicas y –por supuesto– a significar desde múltiples modalidades.

A través de las cifras de este gráfico también se reitera la importante presencia de la **Fórmula** en el ámbito de las Ciencias Básicas; no obstante ello, se registra una ocurrencia porcentual de 3,47% en las Ciencias Sociales y Humanidades, lo cual indica que su empleo, aunque escaso, debe aportar a la construcción de determinados significados textuales multisemióticos. Por su parte tanto el **Gráfico Estadístico** (78,88%) como la **Figura Geométrica** (73,79%) se unen a la **Fórmula** para construir un panorama de los artefactos multisemióticos más característicos del discurso de las Ciencias Básicas, a partir de los datos obtenidos de la muestra de 545 textos.

3.3. Algunos resultados en los textos de las disciplinas

En este apartado revisamos la ocurrencia de los artefactos pero ahora presentados a partir de la ocurrencia en los textos de la muestra en cada una de las seis disciplinas en estudio. El Gráfico 4 presenta información en términos porcentuales de la ocurrencia de los artefactos multisemióticos.

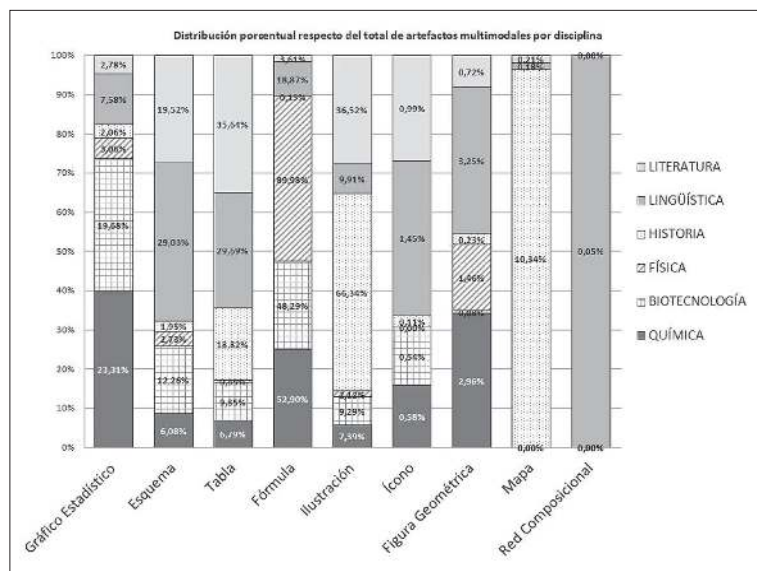


Gráfico 4. Distribución porcentual de artefactos en los textos de cada disciplina.

A la luz de los datos hasta ahora estudiados, se hace muy relevante destacar que la **Fórmula** con su elevada ocurrencia en la muestra se concentra mayoritariamente en la disciplina de Física. El 90,56% del total de las ocurrencias en la muestra de textos se encuentra en los 181 textos de Física. El resto de las ocurrencias se identifican, comparativamente, de modo muy menor en Química, Biotecnología, Lingüística y Literatura. Las cifras del Gráfico 4 también permiten señalar que, en los textos de Física, junto a la altísima ocurrencia de la **Fórmula**, también se detecta una importante aparición de otros cinco artefactos: **Figura Geométrica** (63,45%), **Gráfico Estadístico** (43,03%), **Esquema** (28,55%), **Ilustración** (25,21%) y **Tabla** (7,38%). Todo ello, en su conjunto, hace notoria una caracterización del discurso académico de la Física y su naturaleza altamente multisemiótica, en la que se identifican seis artefactos multisemióticos con elevada ocurrencia en los textos de la muestra en estudio. Esto quiere decir, por una parte, que en la constitución del discurso de la Física se reconocen complejas relaciones intersemióticas entre los sistemas verbal, gráfico, matemático y tipográfico y, por otra, que esta concurrencia de artefactos multisemióticos se registra en alta presencia.

Complementariamente a los datos de Física, a partir del mismo gráfico, se releva que los nueve artefactos multisemióticos identificados solo se registran en los textos de una sola disciplina científica: Lingüística. Aunque con ocurrencia mucho más baja que lo que acontece en Física, en los textos de la muestra de Lingüística se despliega un amplio y heterogéneo repertorio de recursos multisemióticos. En el polo opuesto se ubica Historia, en cuyos textos se identifican tres artefactos multisemióticos (**Mapa**, **Ilustración** y **Tabla**).

Una advertencia es fundamental. Los porcentajes permiten comparar la ocurrencia de los artefactos a través de los textos de las disciplinas, sin embargo ocultan la ocurrencia efectiva en los textos de la muestra. Esto quiere decir que, por ejemplo, en el caso de los textos de Historia, se consigna que el Mapa presenta un 89,63% de ocurrencia y en los textos de Lingüística, la **Red Composicional** registra un 100%. No obstante ello, en Historia se identifican 353 mapas y en los de Lingüística se encontraron 9 redes composicionales. No cabe duda de la eficacia de la estadística descriptiva por medio de porcentajes, pero se hace necesaria también una mirada a las cifras brutas.

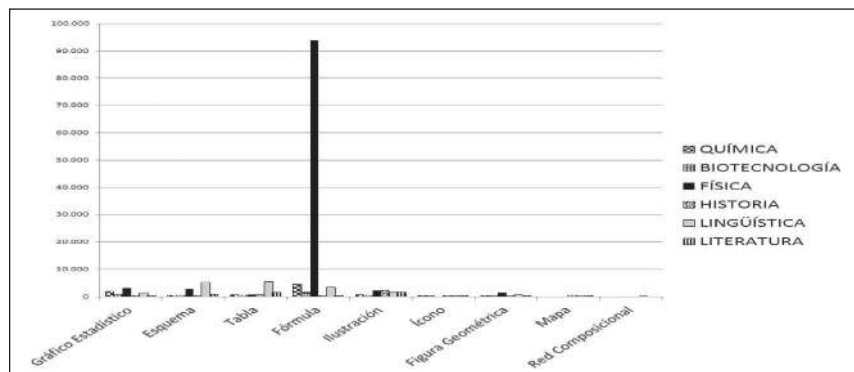


Gráfico 5. Ocurrencia bruta por artefacto y disciplina.

Las cifras entregadas en el Gráfico 5, en parte, reiteran la altísima ocurrencia de la **Fórmula**, en comparación con todos los otros ocho artefactos. Del mismo modo, permiten visualizar comparativamente la diferencia entre la ocurrencia de un artefacto y otro.

3.4. Ocurrencia promedio de artefactos por disciplina y por página

Con el objetivo de obtener un panorama más acotado de la caracterización del discurso multisemiótico disciplinar, procedimos –a partir del total de páginas de los textos de la muestra– a calcular la frecuencia de ocurrencia de cada artefacto por página. Los resultados agrupados por artefacto se presentan en el siguiente gráfico.

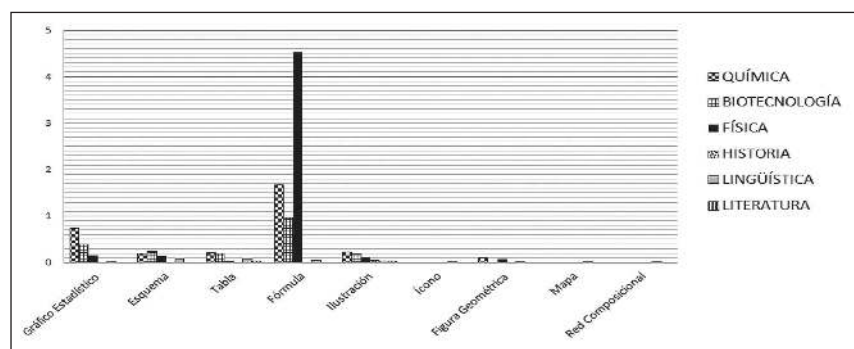


Gráfico 6. Frecuencia de ocurrencia de cada artefacto por página a través de las seis disciplinas.

Estos datos corroboran lo que hemos venido apuntando acerca de la singular constitución del discurso especializado de la Física, pues tal como se aprecia en el

Gráfico 6, una página de un texto de Física presenta una alta ocurrencia promedio de un número importante de artefactos. El rasgo más prominente es, con toda claridad, la alta ocurrencia de fórmulas por página. Situación similar ocurre, proporcionalmente en escala menor, en una página de un texto de Química y en otra de Biotecnología. En comparación, una página de un texto de las disciplinas de las Ciencias Sociales y Humanidades muestra una constitución muy diferente.

Estos datos empíricos comprueban lo que muy posiblemente muchas personas comunes o científicos imaginan de los textos de estas disciplinas. Ciertamente existe un imaginario colectivo acerca del discurso de las Ciencias Básicas en oposición al de las Ciencias Sociales y Humanidades. Lo importante en este caso es su comprobación científica y lo que los hallazgos empíricos nos muestran ciertamente. La corroboración o no de las hipótesis con datos más allá de todo orden meramente especulativo es lo que el rigor científico busca establecer en su ideario. En este caso, asistimos a una corroboración de nuestras intuiciones, es decir, en su fase descriptiva desde la ocurrencia de artefactos multisemióticos, los textos de las Ciencias Básicas presentan una mayor concentración de una amplia variedad de los mismos en promedio en cada página de los textos de la muestra.

Este hallazgo descriptivo no debe confundirse con la idea de que este tipo de discurso por su configuración multisemiótica es necesariamente más complejo de procesar. Es probable que lo sea, pero ésta es justamente otra cuestión que debe investigarse y que la indagación psicolingüística deberá dar cuenta de ello. La vinculación entre una mayor presencia de diversos artefactos multisemióticos en una página de texto y su mayor o menor complejidad de procesamiento y grado de comprensión, memoria y aprendizaje se constituye así en un terreno emergente a partir de los datos aquí entregados.

Otro modo de mirar estos mismos datos es exclusivamente desde las disciplinas y el promedio del conjunto de artefactos que se registran en una página de los textos. El Gráfico 7 muestra estas cifras promedio.

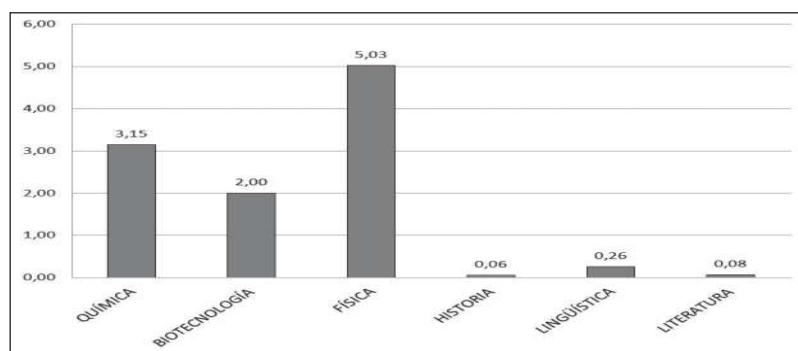


Gráfico 7. Frecuencia de ocurrencia de artefactos por página en cada disciplina.

Estas cifras vuelven a mostrar que la lectura de una página a través de estas seis disciplinas ofrece complejidades de distinto orden en cuanto a la presencia de artefactos multisemióticos. Queda claro que el discurso científico no se constituye únicamente de palabras y variaciones tipográficas, sino que simultáneamente en algunas disciplinas se puede llegar a detectar en promedio más de cinco artefactos por página. Ello implica que en algunos casos es posible encontrar una mucho más alta concentración de estos artefactos pues la normalización de cifras ciertamente no implica homogeneidad total. En efecto, habrá páginas de textos de Física y Química que mostrarán ocurrencias muy superiores a las registradas en el Gráfico 7.

A partir de las cifras de este gráfico también se reitera el llamado de atención hacia la naturaleza altamente diferencial entre las Ciencias Sociales y Humanidades y las Ciencias Básicas. Y, del mismo modo, también es muy singular la variación detectada entre las disciplinas de uno y otro dominio. Queda en evidencia la mayor presencia de artefactos promedio por página en Lingüística (0,26 artefactos en promedio por página), a diferencia de la naturaleza más predominantemente verbal de los textos de Historia (0,06).

Lemke (1998) ya había llamado la atención acerca de la heterogeneidad y alta concentración de artefactos a partir de un análisis preliminar de un grupo de textos de Física y de Medicina en lengua inglesa. Basado en artículos tomados de prestigiosas revistas científicas y algunos capítulos de libros, Lemke (1998) llega de modo preliminar a establecer una ocurrencia promedio de entre 3 a 6 artefactos por página, dependiendo eso sí del género discursivo. Para ello, parte de una clasificación *a priori* y distingue gráficos, ecuaciones y tablas. Interesantemente, aunque con un corpus intencionado en disciplinas y géneros y de tamaño pequeño en comparación con el nuestro, sus resultados apuntan en la misma línea que los nuestros.

4. BREVE DISCUSIÓN: VARIACIÓN MULTISEMIÓTICA DISCIPLINAR

Este acercamiento posiblemente más ecológico, es decir a partir de textos completos y recolectados en contextos reales, en usos auténticos y con metodologías empíricas “basadas en el corpus”, nos permite identificar diversos sistemas de construcción de significados a través de las disciplinas, todos dentro del discurso académico a nivel de doctorado. En esta oportunidad no hemos puesto el acento en el tipo de lengua en uso, es decir, no estamos preocupados de si estos textos están escritos en español o inglés; no obstante ello, cabe dejar constancia que un número considerable de los materiales analizados están originalmente escritos en lengua inglesa, lo cual entrega antecedentes del tipo de lengua que se lee preferentemente a nivel de programas de doctorado.

Ahora bien, los hallazgos aquí reportados permiten distinguir claramente un discurso científico más predominantemente gráfico-matemático (Física, Química y Biotecnología) de otro más sobresalientemente verbal-tipográfico (Historia, Literatura y Lingüística). Esta variación multisemiótica disciplinar se grafica en la siguiente Figura 5.

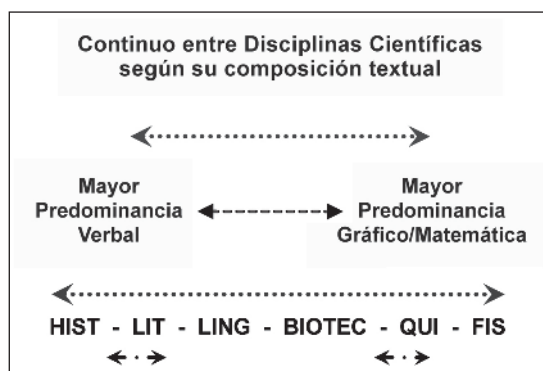


Figura 5. Continuo de disciplinas según constitución textual.

Precisamente, a partir de los datos empíricos recolectados, se reconoce el discurso de la Física y el discurso de la Historia como dos polos dentro de un continuo aún no definitivamente determinado en su constitución interna de textos especializados a través de las disciplinas científicas. Todo ello a partir de la caracterización por identificación en los textos de las disciplinas y por la ocurrencia conjunta de diversos artefactos promedio en una página. Estos hallazgos permiten alinear a los textos de la muestra de estas seis disciplinas científicas en un continuo desde las más predominantemente verbales hasta lo más prototípicamente gráfico-matemático. Según los datos recabados, en una posición intermedia se sitúan los textos de Lingüística y Biotecnología, los cuales interesantemente se acercan en sus rasgos multisemióticos.

Ahora bien, esta constitución textual y su respectiva construcción de significados multisemióticos conducen a una reflexión en torno a la configuración de los espacios textuales disciplinares y la composición de los artefactos mismos. La lengua en su riqueza por brindar a los seres humanos de un potencial de significado ofrece una multiplicidad de recursos en la composición textual y la selección de los mismos se realizan en los textos escritos estáticos a través de las disciplinas tanto en la elaboración de los artefactos multisemióticos como en la construcción definitiva de los textos especializados. En este marco de estudio, lo que está en discusión es si los especialistas de unas y otras disciplinas, en conocimiento de los diversos recursos disponibles, han optado por preferir unos y otros porque ellos dan cuenta de los significados que vehiculan ese tipo

de conocimiento particular o porque ese tipo de conocimiento no puede ser comunicado mediante otros sistemas semióticos. Esto quiere decir que los científicos de la Física deben obligatoriamente hacer uso de los artefactos que han llegado a construir como único medio de significar por escrito sus mensajes y que los científicos de la Historia han preferido significar preferiblemente a través del sistema verbal sus conocimientos porque éste se ajusta a la naturaleza de ese conocimiento. Así, a partir de los hallazgos empíricos aquí presentados, es posible delinear un continuo de variación a partir de los textos de las disciplinas científicas según su composición textual multisemiótica.

5. COMENTARIOS DE CIERRE

Como ha quedado demostrado empíricamente a lo largo de este artículo, la ciencia se comunica por medio de diversos sistemas semióticos, los que se configuran de manera particular en cada disciplina. En este sentido, los textos de la ciencia son irreducible y esencialmente de naturaleza multisemiótica. La construcción de conceptos científicos se realiza por medio de significados interrelacionados con base en una diversidad de complejas modalidades que incluyen –entre otras– a la verbal, la gráfica, la matemática y la tipográfica.

Así, los datos aportados en esta investigación revelan, con base en una muestra de 1.043 textos a través de seis disciplinas científicas, que para hacer y comunicar ciencia, es decir, para construir y transmitir conocimiento científico se combinan diversas sistemas canónicos a partir de los cuales ha sido posible identificar nueve artefactos multisemióticos: **Esquema, Figura Geométrica, Fórmula, Gráfico Estadístico, Ícono, Ilustración, Mapa, Red Composicional y Tabla.**

Ahora bien, la inmersión ecológica y sin restricciones en el mundo del discurso académico a través de un conjunto de disciplinas científicas y sus singularidades ha resultado una experiencia notable. Ello ha implicado –entre otros– la revisión permanente y puesta en discusión del concepto de texto escrito estático y el cuestionamiento a los rasgos constitutivos de su naturaleza misma, así como la indagación crítica del concepto de multimodalidad y el descubrimiento de confusiones no solo terminológicas sino también conceptuales.

Resulta magnífico –a la luz de una investigación empírica– poder comprobar el avance científico de tipo acumulativo y el enriquecimiento de los objetos de estudio en nuestro progresivo caminar hacia el escrutamiento de ciertos fenómenos científicos como es, en este caso, el texto escrito estático. Desde una mirada acentuadamente lexicogramatical y bajo la hegemonía de lo verbal, con el afán de desarrollar etiquetadores morfosintácticos automáticos, no habíamos capturado la tremenda riqueza del discurso académico disciplinar en su naturaleza intrínsecamente multisemiótica.

Los hallazgos empíricos reportados en este estudio, aunque a partir de una muestra aleatoria y proporcional al 33%, brindan un panorama de conjunto del discurso académico especializado en los textos de seis disciplinas científicas del que no se tiene registro en la lengua española ni en otra. La singularidad de este estudio, la robustez empírica de los datos ecológicamente recolectados en contextos de uso cotidiano y en el marco de programas de formación doctoral con acreditación nacional y en universidades públicas y tradicionales de Chile.

Este tipo de hallazgos permiten, en nuestra opinión, formular opiniones informadas adecuadamente, las cuales –entre otras– pueden guiar decisiones de renovación curricular, orientar innovación en políticas educativas, regular la generación de estándares de desempeño y generar lineamientos para planificaciones lingüísticas en contextos diversos. De modo más específico, estos datos también iluminan decisiones curriculares en el marco de la alfabetización académica y profesional a través de las disciplinas y el desarrollo de habilidades de lectura y escritura especializadas. Todo ello en el marco de la construcción, acceso, desarrollo y transmisión de conocimientos disciplinares.

Queda pendiente el estudio de los artefactos multisemióticos a través de las disciplinas, pero por género discursivo. Esta otra mirada aportará información fundamental para conocer aún más de cerca los rasgos del discurso de la ciencia. Nuestra propuesta de cuatro sistemas semióticos: verbal, gráfico, matemático y tipográfico, también deberá ser revisada y posiblemente ampliada a la luz de nuevos datos empíricos.

También se proyecta la necesaria extensión de la identificación de los artefactos en el total de los 3.160 textos del Corpus PUCV-2010, pues este estudio preliminar podría estar solo mostrando una parte de la tremenda variedad de artefactos multisemióticos presentes en los otros textos del corpus.

Por último, un desafío futuro es la indagación en el procesamiento psicolingüístico de estos textos multisemióticos en las diversas disciplinas y su posible grado diferencial de complejidad llegado el momento de construir una representación mental que dé cuenta de los significados construidos a partir de diversos sistemas semióticos. En este momento ya hemos emprendido algunas exploraciones en este terreno, de las cuales esperamos notificar prontamente.

REFERENCIAS

- Alvarado, M. 1994. *Paratexto*. Buenos Aires: Eudeba.
- Baldry, A. 2007. "The role of multimodal concordances in multimodal corpus linguistics". En Royce, T. & W. Bowcher (Eds.) *New directions in the analysis of multimodal discourse*. Mahwah, N.J.: Lawrence Erlbaum, pp. 173-194.
- Baldry, A. & Thibault, P. 2001. "Towards multimodal corpora". En G. Aston &

- L. Burnard (Eds.) *Corpora in the description and the teaching of English*. Bologna: CLUEB, pp. 87-102.
- Bateman, J. 2008. *Multimodality and genre: a foundation for the systematic analysis of multimodal documents*. London: Palgrave Macmillan.
- Bateman, J. 2009. "Discourse across semiotic modes". En J. Renkema (Ed.) *Discourse of course*. Amsterdam: Benjamins, pp. 55-66.
- Bateman, J., Delin, J. & R. Henschel. 2004. "Multimodality and empiricism: Preparing for a corpus-based approach to the study of multimodal meaning-making". En Ventola, E., Cassily, Ch., & M. Kaltenbacher (Eds.) *Perspectives on multimodality*. Amsterdam: Benjamins, pp. 65-90.
- Bednarek, M. & J. R. Martin (Eds.) 2010. *New discourse on language: Functional perspectives on multimodality, identity and affiliation*. London: Continuum.
- Bryan, M. 1988. *SGML: An Author's guide to the standard generalized markup language*. New York: Addison-Wesley.
- CES (Corpus Encoding Standart). 2000. Corpus encoding standart. Version 1.5. <http://www.cs.vassar.edu/CES>.
- Davis, F. 1976. *La comunicación no verbal*. Madrid: Alianza.
- Delin, J., Bateman, J. & Allen, P. 2002/3. "A model of genre in document layout", en *Information Design Journal* 11 (1), pp. 54-66.
- Genette, G. 1997. *Paratexts. Thresholds of interpretation*, Cambridge: CUP.
- Halliday, M. & J. Martin. 1993. *Writing science; Literacy and discursive power*. London: Falmer.
- Jewitt, C. 2009. "Different approaches to multimodality". En C. Jewitt (Ed.) *The Routledge handbook of multimodal analysis*. New York: Routledge, pp. 28-39.
- Kong, K. 2006. "A taxonomy of the discourse relations between words and visual", en *Information Design Journal* 14 (3), pp. 207-230.
- Kress, G. & Van Leeuwen, T. 1992. Structures of visual representations. *Journal of Literary Semantics*, 21(2), 91-117.
- Kress, G. & Van Leeuwen, T. 1996. *A grammar of visual imagery*. London: Routledge.
- Kress, G. & Van Leeuwen, T. 2001. *Multimodal discourse. The modes and media of contemporary communication discourse*. London: Arnold.
- Kress, G. 2009. "What is a mode?" En C. Jewitt (Ed.) *The Routledge handbook of multimodal analysis*. New York: Routledge, pp. 54-67.
- Lemke, J. 1990. *Talking Science: Language, Learning, and Values*. Stamford, Conn.: Ablex Publishing.
- Lemke, J. 1998. "Multiplying Meaning: Visual and verbal semiotics in scientific text". En J.R. Martin & R. Veel (Eds.) *Reading Science*. London: Routledge, pp. 87-113.
- Lemke, J. 2005. "Multimedia genres and transversals", en *Folia Linguistica* 39 (1-2), pp. 45-56.

- Martin, J.R. & Rose, D. 2008. *Genre relations: Mapping culture*. London: Equinox.
- Martin, J.R. & Veel, R. (eds) 1998. *Reading Science*. London: Routledge.
- Martinec, R. & Van Leeuwen, T. 2009. *The language of new media design. Theory and practice*. London: Routledge.
- Matthiessen, C. 2007. "The multimodal page: A systemic functional exploration". En Royce, T. & W. Bowcher (Eds.) *New directions in the analysis of multimodal discourse*. Mahwah, N.J.: Lawrence Erlbaum, pp. 1-65.
- O'Halloran, K. (2004). On the effectiveness of mathematics. En Ventola, E., Casily, Ch., & M. Kaltenbacher (Eds.) *Perspectives on multimodality* (pp. 91-118). Amsterdam: Benjamins.
- O'Halloran, K. 2005. *Mathematical discourse. Language, symbolism and visual images*. London: Continuum.
- O'Halloran, K. 2008. "Systemic functional-multimodal discourse analysis (SFMDA): Constructing ideational meaning using language and visual imagery", en *Visual Communication* 7 (4), pp. 443-475.
- O'Halloran, K., Tan, S., Smith, B. & A. Podlasov. 2010. "Challenges in designing interfaces for the study of multimodal phenomena", en *Information Design Journal*, pp. 18 (1), pp. 2-21.
- Parodi, G. (Ed.). 2005. *Discurso especializado e instituciones formadoras*. Valparaíso: Ediciones Universitarias de Valparaíso.
- Parodi, G. 2006 "El Grial: Interfaz computacional para anotación e interrogación de corpus en español", en *Revista RLA* 44 (2), pp. 91-115.
- Parodi, G. (Ed.). 2007a. *Lingüística de corpus. Puntos de Mira*. Valparaíso: Ediciones Universitarias de Valparaíso.
- Parodi, G. (Ed.). 2007b. *Working with Spanish corpora*. London: Continuum.
- Parodi, G. 2010a. *Lingüística de corpus. De la teoría a la empiria*. Frankfurt: Iberoamericana/Vervuert.
- Parodi, G. (Ed.). 2010b. *Academic and professional discourse genres in Spanish*. Amsterdam: Benjamins.
- Parodi, G. 2010c. "Research challenges for corpus cross-linguistics and multimodal texts", en *Information Design Journal* 18 (1), pp. 69-73.
- Parodi, G. (Ed.). 2010d. *Alfabetización académica y profesional en el siglo XXI. Leer y escribir desde las disciplinas*. Santiago: Ariel.
- Poyatos, F. 1983. *New perspectives in Nonverbal Communication: Studies in Cultural Anthropology, Social Psychology, Linguistics, Literatura and Semiotics*. Oxford: Pergamon Press.
- Poyatos, F. (Ed.). 1992. *Advances in Nonverbal communication*. Amsterdam: Benjamins.
- Royce, T. 2007. "Intersemiotic complementarity: A framework for multimodal discourse analysis". En Royce, T., & W. Bowcher (Eds.) *New directions in the analysis of multimodal discourse*. Mahwah, N.J.: Lawrence Erlbaum, pp. 63-109.

- Saussure, F. 1966. *Course in general linguistics*. New York: McGraw-Hill.
- Scollon, R. & Scollon, S. 2009. "Multimodality and language. A retrospective and prospective view". En C. Jewitt (Ed.) *The Routledge handbook of multimodal analysis*. New York: Routledge, pp. 170-180.
- Stein, P. 2009. "Multimodal instructional practices". En Coiro, J., M. Knobel, C. Lankshear & D. Leu (Eds.). *Handbook of research on new literacies*. New York: Routledge, pp. 871-898.
- Stöckl, H. 2004. "In between modes: Language and image in printed media". En Ventola, E., Cassily, Ch., & M. Kaltenbacher (Eds.) *Perspectives on multimodality*. Amsterdam: Benjamins, pp. 9-30.
- Van Leeuwen, T. 2005. "Typographic meaning", en *Visual Communication* 4, pp. 137-149.
- Van Leeuwen, T. 2006. "Towards a semiotics of typography", en *Information Design Journal* 14 (2), pp. 139-155.
- Wignell, P. 2007. *On the discourse of social science*. Darwin: Charles Darwin University Press.