



Revista Signos

ISSN: 0035-0451

revista.signos@ucv.cl

Pontificia Universidad Católica de Valparaíso  
Chile

Ibáñez, Romualdo

La comprensión del discurso escrito: Una propuesta teórico-metodológica para su evaluación

Revista Signos, vol. 45, núm. 78, marzo, 2012, pp. 20-43

Pontificia Universidad Católica de Valparaíso

Valparaíso, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=157021504002>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto



## **La comprensión del discurso escrito: Una propuesta teórico-metodológica para su evaluación\***

### **Written Discourse Comprehension: A theoretical and methodological framework for its assessment**

**Romualdo Ibáñez**

romualdo.ibanez@ucv.cl

Pontificia Universidad Católica de Valparaíso

Chile

Recibido: 6-I-2011 / Aceptado: 4-VIII-2011

**Resumen:** Durante las últimas décadas, la evaluación de la comprensión del discurso escrito ha concitado el interés de una gran cantidad de especialistas, provenientes de diferentes ámbitos disciplinares, como la psicolingüística, la psicología del discurso, la lingüística aplicada, etc. Esto ha derivado en un número no menor de trabajos investigativos, los que, en gran parte, se centran en diversas problemáticas relacionadas con los instrumentos de evaluación. Tales problemáticas incluyen, entre otras, la pertinencia de las técnicas de evaluación, la incidencia de las características de los textos en la medición, el número y tipo de procedimientos necesarios para la validación de los instrumentos. Sin embargo, a pesar de la gran diversidad de estudios realizados en el área, son escasos los trabajos que, desde una perspectiva teórico-metodológica claramente definida, orientan la generación de instrumentos de medición. En este trabajo, proponemos un marco para el diseño y construcción de instrumentos de evaluación de la comprensión de textos escritos, basado en nuestra concepción teórica del fenómeno. En la propuesta, además, destacamos el valor de los estudios de corpus para sustentar las decisiones del especialista. Ello pues, este tipo de estudios permite guiar el establecimiento y mantención de una relación coherente entre los constructos teóricos a medir, los textos a seleccionar y las técnicas de evaluación a utilizar, lo que, en definitiva, debería redundar en un mejor instrumento de medición.

**Palabras Clave:** Evaluación, comprensión de textos escritos, instrumentos.

**Abstract:** During the last decades, assessment of written discourse comprehension has attracted the interest of a large number of specialists from various disciplinary fields, such as psycholinguistics, discourse psychology, applied linguistics, among others. This has led to a significant number of studies, most of which are essentially focused on issues regarding assessment instruments. Such issues include, among other aspects, how pertinent assessment techniques are, how important text's features are in the measuring, the number and the type of procedures required for validating instruments. However, in spite of the great variety of studies in the area, frameworks to guide the creation of assessment tools from a clear theoretical-methodological perspective are scarce. In this paper, we propose a framework for the design and construction of instruments for assessing written text comprehension, based on our theoretical conception of the phenomenon. Besides, in this proposal, we emphasize the importance of corpus studies to support the specialist's decision making during the design and construction of the instrument. Corpus studies can guide the generation and maintenance of a coherent relation among the theoretical constructs to be measured, the texts to be selected, and the assessing techniques to be used. In turn, this should result in a better assessment instrument.

**Key Words:** Evaluation, written text comprehension, instruments.

## INTRODUCCIÓN

La relevancia asignada actualmente a la comprensión del discurso escrito, no solo por parte de especialistas y autoridades gubernamentales, sino que también por el común de las personas, ha hecho de este un fenómeno psicolingüístico de gran interés científico a nivel mundial. Tal interés se ve reflejado en la gran cantidad de investigaciones al respecto. Particularmente en Chile, la comprensión del discurso escrito ha llamado la atención de muchos investigadores. Ejemplo de ello es el fructífero trabajo realizado por los especialistas de la Pontificia Universidad Católica de Valparaíso y, más precisamente, de la Escuela Lingüística de Valparaíso (ELV), donde se han realizado descripciones cada vez más precisas de los desempeños de los lectores en diferentes contextos de lectura (Parodi, 1991, 2000, 2002, 2007; Peronard, 1999, 2007) y en diferentes lenguas (Ibáñez, 2007, 2008; Santana, 2010). También se ha estudiado la relación entre tal desempeño y otras variables como el uso de estrategias (Peronard, 1989, 1992) y el desarrollo de mecanismos metacognitivos (Peronard, Gómez Macker, Parodi, Núñez & González, 1997; Peronard, Crespo & Velásquez, 2000, 2002; Peronard, Crespo & Guerrero, 2001; Peronard, Velásquez, Ibáñez, Ordenes & Alonzo, 2006). Se han propuesto taxonomías de inferencias (Parodi, 2005) y hasta modelos que describen la comprensión del discurso escrito (Peronard & Gómez Macker, 1985; Parodi, 2011). Lo que sucede en Chile es solo reflejo de la gran relevancia que se le asigna a la comprensión del discurso escrito en todo el mundo.

Indudablemente, en el escenario anteriormente descrito, utilizar procedimientos evaluativos eficaces es fundamental, razón por la cual, un número no menor de especialistas se ha dedicado a desarrollar investigación en el área. Esto ha resultado en una gran cantidad de estudios que abordan diversas materias vinculadas con la evaluación de la comprensión del discurso escrito y, más precisamente, con la generación de instrumentos de medición, como son, entre otras, los enfoques evaluativos (Eeds, 1988), las técnicas de evaluación (Nuttall, 1996; Rauch & Hartig, 2010), las características de los textos (Crossley, Allen & McNamara, 2011), la validez de los instrumentos (Alderson & Banerjee, 2002; Weir, 2005). Asimismo, algunas interrogantes frecuentes entre los especialistas en el área son: ¿Qué relación existe entre las características de los textos y los niveles de desempeño? ¿Existen procesos psicolingüísticos que implican mayor dificultad a la hora de evaluar? ¿Qué técnica de evaluación es más apropiada para medir un determinado constructo teórico?

Las interrogantes anteriormente señaladas dan cuenta del constante interés de los especialistas que se han dedicado a estudiar la evaluación de la comprensión del discurso escrito por mejorar los procedimientos evaluativos y, de esa manera, contribuir a que el propósito establecido para una evaluación sea alcanzado cabalmente por el evaluador. Según Weir (2005), una condición fundamental para alcanzar el o los propósitos de una evaluación es que los instrumentos utilizados otorguen datos que realmente den cuenta de las habilidades que se

desean evaluar. Desafortunadamente, y si bien existe gran cantidad de bibliografía en torno a la evaluación de la comprensión de textos escritos, se aprecia una escasez de propuestas que orienten el diseño y la construcción de instrumentos, sustentadas fuertemente en una teoría de la comprensión del discurso escrito y en metodologías que permitan respaldar las decisiones de los especialistas con datos empíricos.

En este trabajo presentamos una propuesta teórico metodológica para el diseño y construcción de instrumentos de evaluación de la comprensión de textos escritos. La propuesta se sustenta en nuestra concepción teórica del fenómeno y enfatiza la importancia de los estudios de corpus como un medio para guiar una vinculación coherente entre los constructos teóricos a medir, las técnicas de evaluación a utilizar y los textos a seleccionar. Para ello, realizamos una revisión teórico-bibliográfica de ciertos marcos conceptuales centrales para nuestra propuesta. Posteriormente, abordamos la relación entre instrumento y propósito de la evaluación. Por último, proponemos una articulación de orden teórico-metodológico entre el o los constructos teóricos a medir, las técnicas de evaluación a utilizar y los textos a seleccionar, de acuerdo a las características de los géneros discursivos a los cuales estos pertenecen.

## **1. La evaluación y el procesamiento del lenguaje**

Entendemos que para abordar la evaluación de textos escritos es preciso comenzar por explicitar, al menos en términos generales, nuestra concepción de evaluación. Para ello, resulta necesario comenzar por señalar que la evaluación constituye un concepto amplio y, en ocasiones, difícil de delimitar. Tal ambigüedad conceptual obedece no solo a los diferentes ámbitos y situaciones en los que el término es utilizado, sino que también a las diferentes concepciones que históricamente se han manejado, dependiendo de los paradigmas epistemológicos que han determinado el desarrollo de las ciencias (Ahumada, 2002). Cabe precisar, además, que si bien entendemos que la evaluación constituye parte fundamental de diversas ciencias y áreas del saber, en este trabajo, nos focalizamos en la concepción que se tiene desde ámbitos disciplinares pertenecientes a las ciencias sociales y humanas.

De acuerdo al Diccionario de la Real Academia de la Lengua (DRAE, 2011), evaluar significa valorar o

asignar valor a algo. Desde una concepción amplia, entonces, podríamos asumir que evaluar constituye parte fundamental de toda actividad humana, en tanto la valoración es requisito de cualquier toma de decisiones. Sin embargo, en ámbitos especializados, el término evaluación, es decir, la connotación más técnica de la palabra, se refiere a un proceso en el que se obtienen datos, producto de una serie de procedimientos concurrentes, llevados a cabo de manera sistemática con el propósito de tomar decisiones (Alderson, 2000; Ahumada, 2002; Bachman, 2004). Evidentemente, la connotación más técnica difiere del significado más general. No obstante ello, su propósito final sigue siendo el mismo, tomar una decisión. La distinción entre ambas concepciones se basa en un criterio de objetividad. En ambos casos se realiza una valoración para tomar una decisión; en el primero sin, necesariamente, contar con datos empíricos que sustenten el valor asignado y, en el segundo, con el sustento de un proceso sistemático.

La diferencia, en términos de objetividad, que se puede establecer entre la concepción cotidiana de evaluación y la que se maneja en ámbitos de especialidad puede ser atribuida al desarrollo científico histórico que han experimentado disciplinas e interdisciplinas como la psicología, la educación y la psicolingüística. De acuerdo a Ahumada (1983), es posible distinguir, a lo menos, tres momentos en la evolución histórica del concepto de evaluación. En un comienzo, principalmente en contextos educacionales, la evaluación era entendida en términos de juicio de especialista. El principio fundamental se basaba en la experiencia del maestro, la cual lo capacitaba para valorar el desempeño de los estudiantes. La mayor desventaja que presentaba esta forma de ver la evaluación era el alto grado de subjetividad. Posteriormente, a principios del siglo XX y con el propósito de dar mayor objetividad al juicio, se comenzó a entender la evaluación en términos de medición. En esta época son reconocidos los trabajos de Thorndike (1904) y Simon (1957), quienes, desde la psicología y con una fuerte influencia conductista, propusieron mediciones sistemáticas para dar cuenta de los aprendizajes de los individuos. Esta concepción de la evaluación se basaba en un carácter psicométrico, sustentado en la aplicación de cálculos estadísticos, lo que posibilitaba la generalización de las mediciones a comportamientos de grupos mayores. Este nuevo planteamiento superaba las limitaciones en términos de subjetividad que implicaban los juicios personales. Sin embargo, esta forma de entender la evaluación

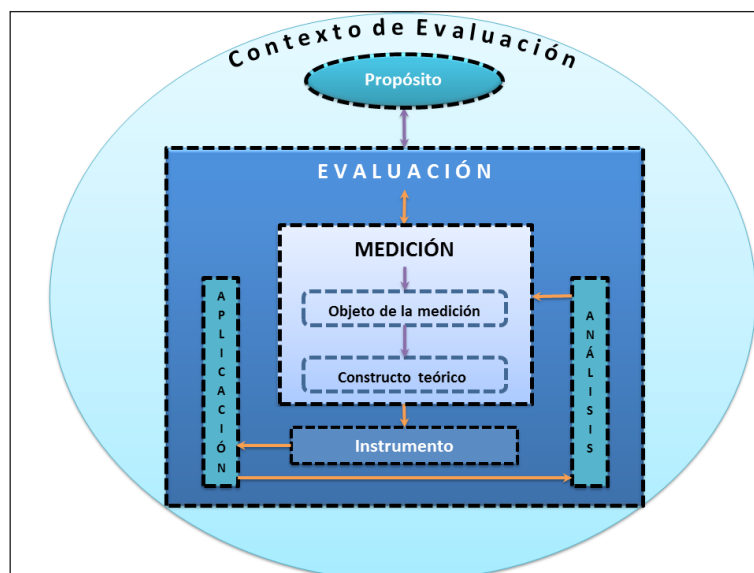
fue criticada pues, la medición comenzó a entenderse como un fin y no como un medio, llevando a medir conductas, en ocasiones, irrelevantes. Así, a mediados del siglo XX, con propósitos principalmente educacionales, surge la denominada evaluación por congruencia, una nueva concepción a partir de los planteamientos de Tyler (1949), quien sostiene que en cualquier proceso evaluativo debe existir coherencia entre la medición y los propósitos de la evaluación. Estos planteamientos llevaron a poner énfasis en los propósitos de la evaluación y en el cuidado de los procesos de enseñanza aprendizaje, prestando especial atención al rol de la retroalimentación. De acuerdo a Ahumada (1983, 2002), la concepción especializada de evaluación que actualmente se maneja, integra las diferentes perspectivas históricas. En ella se enfatiza el juicio objetivo basado en evidencias, en pos de un propósito.

Es preciso señalar que si bien, históricamente, una de las mayores preocupaciones de los especialistas ha sido la búsqueda de objetividad en la evaluación, desde una perspectiva más bien epistemológica, la evaluación no podría ser vista dicotómicamente en términos de objetiva y no objetiva. Ello debido a la naturaleza misma de la relación entre sujeto y objeto de estudio, la cual en ningún caso llegará a ser definitivamente objetiva (Bunge, 2004). La premisa anterior nos lleva a asumir la idea de que solo es posible alcanzar un menor o un mayor grado de objetividad. De acuerdo a Leslie y Caldwell (2009), es precisamente el propósito de alcanzar un mayor grado de objetividad el que conduce a los especialistas en evaluación a proponer procedimientos cada vez

más exhaustivos. Como se aprecia en la Figura 1, este tipo de procedimientos depende de una serie de elementos interrelacionados.

De acuerdo a lo que hemos venido sosteniendo, la evaluación puede ser entendida como un proceso situado y guiado por un propósito, en el cual se recoge información acerca de un determinado objeto de estudio de acuerdo a procedimientos sistemáticos y teóricamente sustentados (Alderson, 2000; Bachman, 2004). Como se observa en la Figura 1, la evaluación da pie a la medición que, en definitiva, corresponde a la cuantificación de las características de un objeto de interés-operacionalizado a partir de un constructo teórico- de acuerdo a reglas y procedimientos específicos. La medición, por su parte, se concreta en instrumentos, los que corresponden a tipos particulares de medición diseñados y contruidos para elicitat conductas específicas. Será a partir de la aplicación de los instrumentos que se obtendrán los datos necesarios que, mediante su análisis, permitirán sustentar la evaluación y, en este sentido, también las decisiones a tomar.

Como ya señaláramos anteriormente, la evaluación es utilizada en diversos ámbitos disciplinares, con diferentes propósitos y orientada a diversos objetos de estudio. Cuando el objeto de estudio es el lenguaje y su procesamiento, la evaluación es utilizada, principalmente, en el ámbito de la educación, la lingüística aplicada y la psicolingüística (Bachman, 2004; Rauch & Hartig, 2010). En el primero de los casos, se utiliza para realizar diagnósticos,



**Figura 1.** Elementos constituyentes de la evaluación.

ubicar a los alumnos de acuerdo a sus capacidades, certificar el dominio de una lengua, etc. (Alderson, 2000). En el caso de la psicolingüística y la lingüística aplicada, esta se utiliza en investigaciones orientadas, por ejemplo, a determinar el estado de desarrollo de ciertas habilidades, a identificar la incidencia de ciertos factores en el procesamiento psicolingüístico o a identificar los factores determinantes en la adquisición de una L2. Cualquiera sea el propósito con el que se realiza un proceso de evaluación, es fundamental que exista coherencia entre tal propósito y los instrumentos utilizados.

### **1.1. Los instrumentos**

Como anteriormente sosteníamos, los instrumentos corresponden a la manifestación concreta de la medición y, en ese sentido, son diseñados y contruidos para elicitat conductas específicas que permitan a los especialistas asignar valores a tales conductas de modo más certero. Por esta misma razón, es necesario que los instrumentos sean efectivos. De acuerdo a nuestra concepción de la evaluación y en concordancia con diversos autores (Anderson, Bachman, Perkins & Cohen, 1991; Alderson, 2000; Golden & Goldman, 2006), para que un instrumento de evaluación sea efectivo, debe cumplir con dos criterios fundamentales: confiabilidad y validez. El primero se refiere a la consistencia de la medición, es decir, a que los resultados de la aplicación del mismo instrumento de evaluación en diferentes instancias y por diversos evaluadores sean similares. Para ello, se requiere que los evaluadores acuerden los criterios que se emplearán y que se evalúe el mismo desempeño de la misma manera en diferentes oportunidades (Bachman, 2004). Por otra parte, la validez, según Bachman y Palmer (1996), se refiere a la adecuación del instrumento respecto de lo que se desea medir.

Si bien profundizar en aspectos psicométricos como los conceptos de confiabilidad y validez, así como también en los procedimientos implicados en la validación de los instrumentos, está más allá de los objetivos de este trabajo, consideramos pertinente abordar, al menos en términos generales, el concepto de validez, puesto que nuestra propuesta está directamente relacionada con la validez de los instrumentos. En este sentido, es preciso comenzar por aclarar que la validez ha sido ampliamente estudiada y, por lo mismo, entendida de diferentes formas. Para Kelly (1927), la validez correspondía a un atributo de los instrumentos, el cual permitía distinguir si el instrumento realmente medía lo que pretendía medir. De acuerdo a Cronbach (1971), el

especialista no valida un instrumento, más que eso, lo que hace es validar la interpretación de los datos obtenidos en un determinado proceso evaluativo. Según Alderson y Banerjee (2002), esta concepción de la validez trae consigo la idea de que el proceso de validación no concluye al momento de finalizar la construcción de los instrumentos, sino que debe ser entendido como un proceso constante de monitoreo y actualización de información.

Otro punto a considerar tiene que ver con los tipos de validez. Si bien, se han distinguido diferentes tipos de validez, como la validez de contenido, la validez predictiva o la validez de constructo (Hatch & Lazaraton, 1991; Chapelle, 1999), también se sostiene que ninguno de estos tipos de validez es suficiente por sí solo (Shepard, 1993). Sobre este punto, diversos especialistas (Messick, 1989; Bachman, 1990) han propuesto que la validez constituye un concepto unificado. Para Messick (1994), por otra parte, el concepto de validez está centrado en la validez de constructo, cuyo valor no se limita a la interpretación de los resultados obtenidos, sino que debe ampliarse a la utilidad y consecuencias de estos. Esta concepción de la validez destaca la importancia del constructo teórico y su definición en el diseño de los instrumentos. Además, pone de manifiesto que la evaluación no implica solamente conocimientos psicométricos y habilidades para el análisis de los datos, sino que también un conocimiento acabado y actualizado de la teoría por parte del evaluador.

De acuerdo a lo anteriormente planteado, garantizar la validez de los instrumentos durante la etapa de diseño, implica principalmente definir con claridad los constructos a evaluar. De acuerdo a Bachman (2004), un constructo corresponde a una conceptualización teórica de un fenómeno. Debido a que los estudios realizados en torno al procesamiento del lenguaje involucran fenómenos no observables directamente, solo es posible acceder a ellos a partir de la interpretación de ciertas conductas. Para garantizar que la interpretación de la conducta realmente dé cuenta de lo que se desea evaluar, es preciso definir los constructos en términos conceptuales y en términos operacionales. La definición conceptual de un constructo debe sustentarse teóricamente en términos precisos, de modo que a la hora de interpretar las conductas no haya confusión respecto de lo que se va a medir. A modo de ejemplo, si se desea medir el constructo de Comprensión, no es lo mismo definirlo conceptualmente de acuerdo a la propuesta de Kintsch y van Dijk (1978) que hacerlo en base a la propuesta de van Dijk y Kintsch (1983). La definición operacional del constructo, por otra

parte, implica la especificación de los procedimientos y condiciones por medio de los cuales se realizarán las mediciones que permitirán realizar inferencias respecto del fenómeno que se desea medir. Esta definición incluye una especificación de las tareas implicadas en la medición. A modo de ejemplo, si se desea medir el constructo de Macroestructura, este puede ser definido operacionalmente en términos de: escribir un texto de trescientas palabras, que dé cuenta de las ideas principales del texto leído.

Asumiendo la relevancia que el dominio teórico tiene en el diseño de los instrumentos evaluativos y en virtud del objetivo de este trabajo, a continuación presentamos la concepción del proceso de comprensión de textos escritos que guiará nuestra propuesta teórico-metodológica para el diseño de los instrumentos.

## **2. Algunos constructos teóricos implicados en la comprensión del discurso escrito**

La comprensión del discurso escrito es, para nosotros, un fenómeno psicodiscursivo, de naturaleza tanto social como cognitiva. En su dimensión social, la comprensión constituye una actividad comunicativa situada. Más concretamente, nos referimos a que cada vez que un individuo enfrenta un texto lo hace en una situación comunicativa particular y, por ende, con ciertas metas que establecerá a partir de las necesidades que le imponga el contexto. De ahí, la relación incuestionable que se produce entre lector como ente social, el texto como instancia de lengua y las metas de lectura que el lector establece. En cuanto a su naturaleza cognitiva, la comprensión del discurso constituye un proceso intencionado de alta complejidad, conformado por una serie de procesos psicolingüísticos, sustentados a su vez, en una variedad de procesamiento cognitivos de orden inferior (atención, percepción y memoria) y de orden superior (toma de decisiones, monitoreo, reflexión, entre otros). Así, a partir de la interacción de la totalidad de los procesos involucrados, se genera una representación mental de la situación descrita en el texto, en base a la información textual, a las metas del lector y a sus conocimientos previos (Parodi, Peronard & Ibáñez, 2010).

Debido a los objetivos de este trabajo, y sin ignorar que la comprensión involucra tanto oralidad como escritura, nuestro interés solo se focaliza en los procesos involucrados en la comprensión del discurso escrito. Así las cosas, y en concordancia con planteamientos presentados en propuestas anteriores (van Dijk & Kintsch, 1983; Kintsch, 1988,

1998, 2002; Graesser, Singer & Trabasso, 1994), es posible sostener que en la comprensión profunda de un texto escrito, los procesos psicolingüísticos involucrados interactúan para la generación de lo que se podría denominar una representación de tres niveles de profundidad, equivalentes al Código de Superficie, la Base Textual y al Modelo de Situación; y dos planos estructurales, constituidos por la Micro y la Macroestructura.

Cabe señalar, además, que la concepción adoptada, así como los constructos correspondientes a cada uno de los niveles y planos, solo responde a una segmentación de orden metodológico que permite a los especialistas describir de mejor manera el fenómeno en investigación. En este mismo sentido, procederemos a referirnos brevemente a cada uno de estos constructos.

### **2.1. Los niveles de representación del discurso**

El Código de Superficie corresponde a la representación de palabras y frases del texto. Este nivel de representación del discurso, está directamente ligado a los procesos automáticos de decodificación, entre los que es posible identificar la decodificación de los símbolos gráficos y la decodificación sintáctica. Aquí, la percepción, la atención y la memoria de corto plazo juegan un rol central en el reconocimiento y asociación de sílabas y palabras. Si bien, nuestra perspectiva no se focaliza en los procesos de decodificación, no se debe restar importancia al rol que estos tienen en lo que se entiende por comprensión profunda de un texto escrito, en cuanto resultan fundamentales para la generación de la representación global de la situación descrita en el texto (Perfetti, Yang & Schmalhofer, 2008).

La Base Textual constituye una representación mental del significado del texto, al que diversos autores (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983; Kintsch, 1988) asignan un formato representacional de orden proposicional. Los textos, generalmente describen situaciones, ya sea reales o imaginarias, las cuales no pueden ser comprendidas cabalmente si el lector no integra la Base Textual que ha construido de un texto determinado a su conocimiento previo. La construcción de un Modelo de Situación, que consiste, básicamente, en la vinculación de la representación semántica del texto con los conocimientos almacenados en la memoria episódica permite llegar al nivel más profundo de comprensión del texto leído. Esto pues se asume que tal integración produce una modificación en

la estructuración de los conocimientos previos del lector. Así, la gran diferencia entre alcanzar un nivel correspondiente a la Base Textual y uno correspondiente al Modelo de Situación no radica en la cantidad o tipo de conocimiento previo utilizado para la construcción de la representación, sino que, precisamente, en la integración de esta nueva información a los conocimientos almacenados en la memoria del lector (Kintsch, 1988, 1998; Mannes & St. George, 1996; McNamara & Kintsch, 1996; McNamara, 2004). De este modo, llegar a construir un Modelo de Situación permite al lector no solo comunicar lo que ha leído (Peronard & Gómez Macker, 1985), sino también aplicar el nuevo conocimiento adquirido a través de su lectura a nuevas situaciones (Kintsch, 1998). En otras palabras, comprender profundamente un texto, se asocia con el aprendizaje. Ahora bien, la Base Textual y el Modelo de Situación no deben ser entendidos como dos niveles absolutamente diferentes y desvinculados. De hecho, corresponden a representaciones en muchos aspectos interdependientes.

## 2.2. Los planos estructurales

Directamente vinculados con los niveles de representación del discurso se encuentran los constructos de Microestructura y Macroestructura (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983). La vinculación podría ser entendida a modo de mutua dependencia en cuanto alcanzar el nivel de la Base Textual requiere de la generación constante de una Microestructura, mientras que la Macroestructura solo puede construirse habiendo alcanzado el nivel correspondiente al Modelo de Situación.

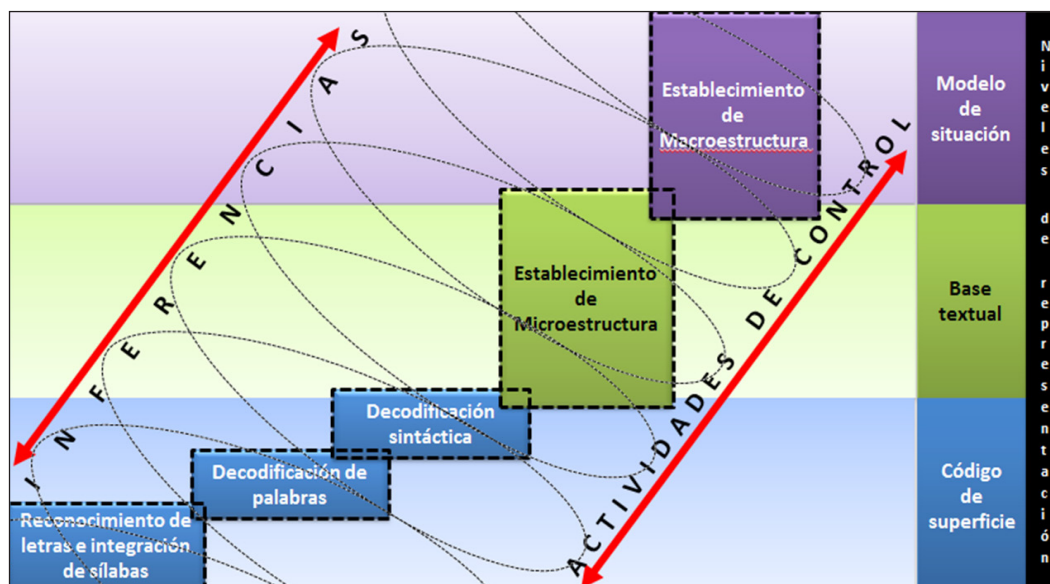
La construcción de la Microestructura se genera a partir de procesos de orden semántico que interactúan directamente con los procesos de decodificación, pues su construcción consiste en la formación de una cadena proposicional a partir del establecimiento de las diversas relaciones semánticas explícitas e implícitas que el lector es capaz de representar a partir de la información textual. A este procesamiento, diversos autores (Sanders & Sporeen, 2001; Louwerse, 2002, 2004; Louwerse & Graesser, 2006; Sanders & Pander, 2006) lo han denominado el establecimiento de la Coherencia Local. En esta parte del proceso de comprensión, especialmente cuando se trata de casos en que existe información no explícita en el texto, se considera fundamental llevar a cabo Inferencias de tipo puente (McNamara & Kintsch, 1996; Graesser & Wiemer-Hastings, 1999; McNamara, 2004).

Además de la construcción de la Microestructura, existe un proceso a partir del cual se establece una organización general, dependiente de la interrelación de los tópicos centrales del texto, que será denominada Macroestructura (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983; Kintsch, 1988, 2002; Louwerse & van Peer, 2002; Kintsch & Rawson, 2005). La construcción de esta estructura representacional global del texto involucra la eliminación, generalización y construcción de nueva información a partir de las proposiciones componentes de la Microestructura (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983; Kintsch, 1988, 1998, 2001, 2002). Esta nueva organización de la información nuclear se realiza generalmente en base a los esquemas retóricos asociados a los géneros discursivos, que son parte del conocimiento previo del lector (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983; Kintsch, 1988; McNamara & Kintsch, 1996).

Así también, existen ciertas actividades de control que son fundamentales durante la totalidad del proceso para un resultado exitoso, entre las que se pueden mencionar, la planificación y ajuste de acuerdo al objetivo, el uso del contexto en forma adecuada, y el monitoreo constante del proceso. Además, cabe señalar que gran parte de la comprensión, como un proceso de construcción de significado, está guiado por la generación de diversos tipos de inferencias (Graesser et al., 1994; Graesser, Wiemer-Hastings & Wiemer-Hastings, 2001; León, Otero & Graesser, 2002; Parodi, 2007), tanto a nivel local como global.

Mediante la descripción del proceso de comprensión de textos escritos, hemos pretendido mostrar los principales constructos teóricos implicados, así como también la relación de interdependencia entre ellos. A continuación, a partir la Figura 2, nos interesa graficar la jerarquía, en términos de complejidad, asociada a los procesamiento psicolingüísticos requeridos para alcanzar los diferentes niveles de representación.

La forma de graficar el proceso que presentamos en la Figura 2, de ninguna manera debe llevar a entenderlo como secuencial, sino que definitivamente como uno de tipo simultáneo y recursivo. De este modo, y como se puede observar, para alcanzar los niveles de representación y para establecer los dos tipos de estructura señalados, es necesario llevar a cabo ciertos procesos psicolingüísticos indispensables. Asociados al primer nivel de representación, ubicamos, entre otros procesamiento, el

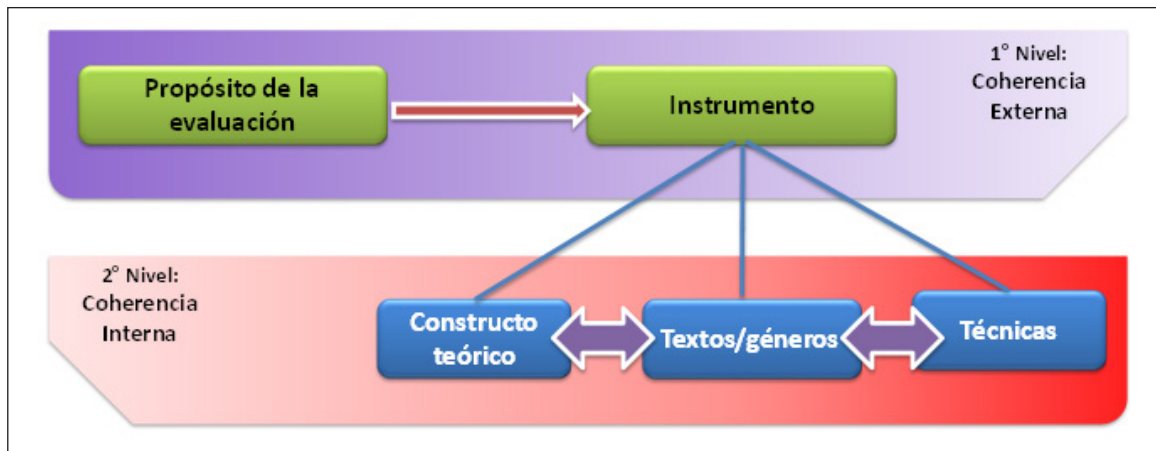


**Figura 2.** Constructos y jerarquía de los procesos psicolingüísticos.

reconocimiento de letras e integración silábica, así como también, la decodificación sintáctica. De acuerdo a nuestra concepción de la comprensión, el grado de dificultad que representan estos procesos es el menor dentro de la jerarquía de dificultad, puesto que a partir de ellos, solo es posible acceder al nivel de representación discursiva correspondiente al Código de Superficie. Por otro lado, la generación de inferencias de tipo puente implica un mayor grado de exigencia, en tanto contribuyen a la generación de la Microestructura a partir de la vinculación de información explícita y no explícita en el texto. Este tipo de procesos permite alcanzar el nivel de la Base Textual. En un grado de dificultad mayor situamos la generación de la Macroestructura, ya que implica la articulación de la información, a partir de estrategias de integración, generalización o eliminación. Estos procesos se encuentran al nivel del Modelo de Situación. Como ya señaláramos anteriormente, este último nivel de representación del discurso es considerado el más profundo en cuanto a los niveles de comprensión que se pueden alcanzar, ya que implica la integración del significado del texto con los conocimientos previos, generando una representación de la situación descrita en el texto. Alcanzar este nivel de comprensión implica aprendizaje y, por lo tanto, la capacidad de utilizar dicho conocimiento nuevo en situaciones novedosas y diversas.

### 3. Un marco para la evaluación de la comprensión de textos escritos

Hasta aquí, nos hemos referido a la evaluación y a los principales constructos que componen nuestra concepción de la comprensión del discurso escrito, describiendo además la jerarquía en cuanto a la dificultad implicada en los diferentes procesamiento. A continuación ofrecemos nuestra propuesta para el diseño y construcción de un instrumento de evaluación de la comprensión de textos escritos. Para ello, en primer lugar, presentamos la relación de dependencia entre el instrumento y el propósito de la evaluación, la cual denominamos Coherencia Externa. Posteriormente, en el marco de la Coherencia Interna, nos referimos en detalle a las técnicas de evaluación y a los textos, de acuerdo a sus características como instancias de lengua caracterizables en términos de géneros discursivos. Luego, abordamos la vinculación coherente que hemos venido proponiendo, entre el constructo teórico a medir, las técnicas de evaluación a utilizar y los textos a seleccionar. Por último, presentamos algunos ejemplos que permiten graficar la vinculación anteriormente señalada, la cual consideramos indispensable para generar un instrumento eficaz, que permita alcanzar el propósito de la evaluación.



**Figura 3.** Coherencia Externa y Coherencia Interna.

### 3.1. Coherencia Externa y Coherencia Interna

Las relaciones de coherencia que proponemos pueden entenderse como el vínculo que se debe establecer entre el instrumento y el medio externo y, al mismo tiempo, entre el instrumento y su configuración interna. Como se aprecia en la Figura 3, estos dos tipos de coherencia ocurren en lo que podríamos entender como dos niveles diferentes, pero están estrechamente vinculados y poseen un carácter de interdependencia.

Para graficar la relación de dependencia entre el propósito de la evaluación y el instrumento, resulta pertinente insistir en la idea de que este último solo cumple su objetivo, si permite obtener datos que realmente den cuenta de las habilidades que se desean evaluar. Para ello, es necesario que el instrumento capture y refleje el propósito de la evaluación, por lo que siempre tanto el diseño como la construcción de un instrumento dependerán de tal propósito. A esta relación de dependencia y como se aprecia en la Figura 3, la denominamos Coherencia Externa y la situamos en un primer nivel relacional. Para asegurar la Coherencia Externa, también es preciso que exista una vinculación coherente entre el o los constructos teóricos, los textos y las técnicas de evaluación. A esta vinculación la denominamos Coherencia Interna y la situamos en un segundo nivel relacional.

#### 3.1.1. Relación entre propósito de la evaluación e instrumento: Coherencia Externa

Como diversos autores sostienen (Nuttall, 1996; Bachman, 2004; Pérez Zorrilla, 2005), un aspecto fundamental al momento de llevar a cabo un

proceso de evaluación de la comprensión de textos escritos, cualquiera sea el ámbito en el que tenga lugar, es el propósito con el que este se llevará a cabo. Esto, pues, como se observa en la Figura 3, existe una relación de dependencia entre para qué se desea evaluar y lo que se va a medir. Esta relación de dependencia entre el propósito de la evaluación y el instrumento a utilizar, la que entendemos como Coherencia Externa, se origina debido a que la evaluación constituye un procedimiento de carácter situado, cuyos propósitos están determinados por necesidades contextuales de diferente índole, las cuales pueden tener su origen en contextos como los investigativos, los educativos, los gubernamentales, los institucionales, etc. Así, tales propósitos determinarán el objeto de la medición y, en este mismo sentido, la forma en que este será concebido al diseñar y construir el instrumento.

Cuando los propósitos de la evaluación involucren a la comprensión de textos escritos, los objetos de las mediciones pueden incluir diferentes tipos de procesamiento y/o conocimiento, entre los que es posible mencionar, el reconocimiento de información literal, la generación de inferencias implicadas en el establecimiento de coherencia local o global, el uso estratégico de conocimiento metacognitivo, etc. Para desarrollar la medición, y de este modo alcanzar el propósito de evaluación que se ha establecido, será necesario que, durante el diseño del instrumento, el objeto de medición sea especificado en términos de un constructo teórico, el cual debe, a la vez, ser definido en términos conceptuales y operacionales. De ahí que, por ejemplo, si, en un contexto educativo, la evaluación tiene por propósito ubicar a los alumnos de un curso en diferentes grupos, de acuerdo a sus

capacidades inferenciales, el objeto de la medición estará constituido por los procesos inferenciales. Ello implica identificar y definir conceptual y operacionalmente el número y tipo de inferencias a medir. Asimismo, y dependiendo del propósito de la evaluación, un instrumento puede incluir uno o varios constructos, dependiendo de si se pretende evaluar el proceso de comprensión como un todo o de manera segmentada (Nuttall, 1996).

En este mismo punto, la comprensión puede ser asumida como un proceso o como un producto, dependiendo del propósito de la evaluación.

Cuando se concibe en términos de proceso, las mediciones se llevan a cabo, generalmente, durante la lectura de manera *on line*. En este caso, por ejemplo si se desea medir los procesos de memoria involucrados en la generación del Modelo de Situación, mientras el sujeto lee, se pueden utilizar mediciones aplicables durante el procesamiento (O'Brien, Rizzella, Albrecht & Halleran, 1998; Zwaan & Madden, 2004; O'Brien, Cook & Guéraud, 2010).

Entre ellas se pueden encontrar la medición de los tiempos de lectura, la latencia de respuesta a las pruebas, los protocolos en voz alta (Magliano & Graesser, 1991; McKoon & Ratcliff, 1992; Magliano, Millis, Ozuru & McNamara, 2007). También, la medición se desarrolla en términos de producto, es decir de manera *off line*, cuando el proceso de lectura ha concluido. Este tipo de mediciones buscan indagar, entre otras cosas, en la representación que el lector ha logrado del texto leído (Parodi, 2007; Peronard, 2007; Ibáñez, 2008). Aquí se puede incluir diferentes técnicas de evaluación como las preguntas abiertas, las de selección múltiple, el resumen, etc.

Cabe mencionar, además, que, dependiendo de los propósitos de la evaluación, también se puede usar un método concurrente, en el cual se integren mediciones de tipo *on line* y *off line* (Lomicka, 1998). Así también, las mediciones de tipo *on line* pueden ser utilizadas para dar cuenta de la comprensión entendida como producto.

### 3.1.2. Relación entre constructo teórico, técnicas de evaluación y textos: Coherencia Interna

Como señalamos anteriormente, el instrumento representa la materialización del propósito de la evaluación. En este sentido, este debe medir de la mejor manera el constructo teórico definido. Para eso, en el caso de la evaluación del texto

escrito, es necesario tomar decisiones adecuadas tanto respecto de las técnicas de evaluación a utilizar (Pérez Zorrilla, 2005), como de los textos a seleccionar (Nuttall, 1996). Esto es lo que denominamos Coherencia Interna. Considerando que ya abordamos el concepto de constructo teórico, a continuación, nos referimos a las técnicas de evaluación y a los textos.

#### 3.1.2.1. Las técnicas de evaluación

En términos generales, las técnicas o métodos de evaluación (Alderson, 2000) corresponden a los medios a partir de los cuales se elicitán las conductas que se desean medir y son, generalmente, planteadas en términos de preguntas o de ítems de completación. La gran variedad de técnicas utilizadas en la evaluación de la comprensión de textos escritos puede ser clasificada en dos grupos, las de tipo abierto y las de tipo cerrado. El primer tipo incluye aquellas técnicas que implican la producción, ya sea escrita u oral por parte del sujeto en evaluación, mientras que el segundo incluye aquellas que implican la elección desde un conjunto de alternativas.

A continuación, y a partir de una recopilación de propuestas presentadas por diferentes autores, en la Tabla I, ofrecemos una serie de técnicas utilizadas en la evaluación de la comprensión de textos escritos, junto con sus respectivas descripciones y ejemplos.

Las técnicas de evaluación de la comprensión representan un aspecto fundamental de cualquier proceso evaluativo. Su relevancia es tal que ha dado origen a una gran cantidad de estudios al respecto (Heaton, 1988; Hughes, 1989; Weir, 1990; Paris & Paris, 2001; Afflerbach, 2004; Carlisle & Rice, 2004).

Hasta ahora, mucho se ha debatido respecto de la eficacia y pertinencia de uno u otro tipo de técnicas. Fredrickson (1984) afirma que una evaluación adecuada no puede estar sustentada solo en técnicas de carácter cerrado. Por otra parte, Bennett (1993) señala que aún no se ha probado empíricamente que las técnicas de carácter abierto sean mejores que las de carácter cerrado, producto de su capacidad para medir procesamientos más diversos. Rupp, Ferne y Choi (2006), por su parte, sostienen que las técnicas de tipo cerrado, como la selección múltiple, solo permite aproximarse a procesos psicolingüísticos superficiales. Entre las críticas más frecuentes hechas a las técnicas de carácter cerrado

**Tabla 1.** Técnicas de evaluación.

| <b>Técnica de Evaluación</b> | <b>Descripción</b>  | <b>Ejemplo</b>  |
|------------------------------|---|---|
| • <b>Cloze</b>               | Se presenta un texto con espacios en blanco para su completación. Esta técnica se basa en una distribución aleatoria de las palabras a completar. Más precisamente, se determina un número fijo de palabras entre cada espacio en blanco. Generalmente, es utilizada para medir coherencia global (Alderson, 2000).                                       | <b>Complete el texto que se presenta a continuación.</b><br><br>Amo a mi madre, ____ quiero mucho; pero ella ____ una vida absurda, siempre ____ de un lado a ____ con ese literato, constantemente ____ nombre figura en los periódicos, ____ esto me cansa.   |
| • <b>Cloze Intencionado</b>  | Se presenta un texto con espacios en blanco para su completación. Esta técnica se basa en una distribución intencionada de los espacios a completar. Puede ser utilizada para medir el establecimiento de la coherencia local (Nuttall, 1996).  | <b>Complete el texto que se presenta a continuación sin alterar su sentido.</b><br><br>Usted siempre está filosofando o hablando de dinero. Para usted no hay desgracia mayor que la de ser pobre; ____, para mí es mil veces preferible ir harapiento y pedir limosna...   |
| • <b>Cloze Invertido</b>     | Se presenta un texto con palabras de más. La tarea del lector consiste en eliminar las palabras que no contribuyen al sentido del texto. Esta técnica se utiliza generalmente para medir Macroestructura (Nuttall, 1996).   | <b>En el texto a continuación, elimine las palabras que están de más.</b><br><br>En el prado ya no se despiertan las grullas entonces con su grito ni se oye el zumbir de los moscardones desde de mayo entre el follaje de los tilos.  |
| • <b>Selección Múltiple</b>  | Se presenta una serie de alternativas, entre las cuales se encuentra la o las correctas. Esta técnica puede ser utilizada para medir diversos constructos, tales como coherencia local, coherencia global, Modelo de Situación, etc. (O'Reilly, Sabatini, Bruce & Sukkarieh, 2011).   | <b>Elija la respuesta más adecuada para completar el texto a continuación.</b><br><br>Sí, cada vez me convengo más de que la cuestión no está en las formas viejas o nuevas, sino en que el hombre escriba sin pensar en forma alguna, en que escriba, ____ lo que escribe fluye libremente de su alma.<br><br>a) porque<br>b) aunque<br>c) luego |
| • <b>Verdadero y Falso</b>   | Se presenta una afirmación respecto de cierta información contenida en el texto leído a partir de la cual se debe decidir si es verdadera o falsa. Esta técnica se utiliza generalmente para medir la identificación de información literal; sin embargo, también puede ser utilizada para medir una gran variedad de constructos (Pérez Zorrilla, 2005). | <b>Decida si la afirmación es verdadera o falsa.</b><br><br>De acuerdo al texto leído, el personaje "Shamráiev" tiene una percepción positiva de Pável Semiónich.<br><br>V ____<br>F ____   |

|  |  |   |
|--|--|---|
| • <b>Respuesta Abierta Corta</b>               | Se presenta una pregunta cuya respuesta puede ser sí o no. Esta técnica se utiliza generalmente para medir la identificación de información literal (Alderson, 2000).  | ¿Tiene Shamráiev una buena imagen de Pável Semiónich?   |
| • <b>Respuesta Abierta Larga</b>               | Se presenta una pregunta cuya respuesta debe proveer información específica o general obtenida a partir de la lectura. Esta técnica puede ser utilizada para medir diversos constructos, tales como coherencia local, coherencia global, Modelo de Situación, etc. (Parodi, 2007).   | ¿Cuál es la razón que Juan Pablo Castel usa para escribir su relato?<br>_____<br>_____<br>_____   |
| • <b>Transferencia de Registros Semióticos</b> | Se solicita expresar lo comprendido a través de esquemas, tablas o dibujos. Esta técnica se utiliza generalmente para medir Macroestructura y Modelo de Situación (Nuttall, 1996).   | Represente el Acto Primero de la obra "Gaviota" a través de 1 dibujo.   |
| • <b>Recuerdo Libre</b>                        | Se solicita que, a partir de la lectura, se realice un relato oral que dé cuenta de lo comprendido, sin interrupciones por parte del evaluador y en un tiempo determinado. Esta técnica puede ser utilizada para medir diversos constructos, dependiendo de los objetivos del evaluador. Entre los constructos que se pueden medir se encuentran la microestructura, la Macroestructura, el Modelo de Situación, etc. (Nuttall, 1996). | En un máximo de 5 minutos, cuénteme todo lo que recuerda del texto que acaba de leer.   |
| • <b>Resumen</b>                               | Se solicita la producción de un texto breve que dé cuenta de las ideas principales del texto leído. Esta técnica se utiliza, generalmente, para medir la generación de la Macroestructura (Parodi, 2007).  | A partir del texto leído realice un resumen que contenga no más de 10 líneas.<br>_____<br>_____<br>_____  |
| • <b>Resumen Incompleto</b>                    | Se presenta un resumen sin la totalidad de las ideas que permiten dar cuenta del significado global del texto fuente. La tarea del lector consiste en incorporar las ideas faltantes. Esta técnica se utiliza generalmente para medir la generación de la Macroestructura (Walter, 2004).  | <b>A partir del texto leído, complete el siguiente resumen, incorporando las ideas que faltan.</b><br><br>La integración económica se consigue con la formación de la Unión Aduanera en 1968. _____<br>_____<br>_____. Aún faltaba conseguir la libre circulación de mercancías no industriales; contratar servicios, instalarse y trabajar allí y que el dinero circulase sin restricciones. |
| • <b>Protocolos en Voz Alta</b>                | Se solicita que, a partir de la lectura, se realice un relato oral que dé cuenta de lo comprendido o de la forma en que el lector llegó a comprender lo que leyó. Este relato es guiado por el evaluador a partir de preguntas. Se utiliza principalmente para medir actividades y conocimientos de orden metacognitivo (Magliano & Millis, 2003).   | ¿Qué indicios te ayudaron a comprender la situación en que se encuentra el protagonista al iniciar el relato?<br>¿En qué cosas te fijaste para afirmar que el protagonista sufre un cambio en su personalidad a lo largo del relato?<br>¿Cómo te diste cuenta de que el personaje vivía en un mundo onírico?  |

se encuentra el que a partir de ellas es imposible distinguir si la respuesta del sujeto evaluado fue producto de la comprensión del texto o de otros procesos cognitivos como lo es la resolución de problemas o, incluso, del azar (Alderson, 2000). No obstante ello, diversos autores (Magliano et al., 2007; O'Reilly et al., 2011) resaltan su grado de objetividad y pertinencia.

Desde nuestra forma de entender el asunto, no existe una técnica de evaluación mejor que otra. Su utilidad en un proceso de evaluación estará determinada por el grado de pertinencia con que es utilizada para medir un constructo teórico particular. Así, será tarea y desafío del evaluador seleccionar la técnica más apropiada al momento de diseñar y construir un instrumento, en virtud del constructo a medir y de los textos a utilizar.

### 3.1.2.2. *Los textos y los géneros discursivos*

Se podría decir que las características de los textos han sido foco de interés para los especialistas en evaluación de la comprensión del discurso escrito desde hace bastante tiempo. Ya en la década del cuarenta, Rudolf Flesch publicaba su tesis doctoral titulada *Marcas de un Estilo de Lecturabilidad* (Klare, 1974-1975). En su propuesta, Flesch, utilizando un criterio de cantidad y por medio de una fórmula matemática, pretendía categorizar los textos, asumiendo que los que presentaban mayor número de sílabas por palabra y mayor número de palabras por oraciones serían más difíciles de leer. Con el paso del tiempo, y el desarrollo tanto de nuevas tecnologías como de nuevas concepciones respecto del proceso de comprensión, los especialistas han buscado nuevas y diversas formas de describir los textos y, de este modo, clasificarlos en términos de su grado de complejidad. Así, desde una perspectiva más bien semántica, enfocada en las relaciones interclausulares, se han desarrollado investigaciones orientadas a determinar la medida en que el grado de cohesión de los textos incide en el nivel de comprensión alcanzado. Para ello, McNamara y Kintsch (1996) distinguen textos pertenecientes al ámbito científico como más y menos cohesionados. A partir de estos textos elaboraron instrumentos con el propósito de medir el desempeño de los lectores. Los resultados demostraron que los lectores más hábiles y con mayor conocimiento del tema obtenían mejores logros en los textos menos cohesionados. Por otra parte, los lectores

menos hábiles y con menor conocimiento del tema obtenían mejores resultados con los textos más cohesionados.

Por otra parte, el uso de herramientas computacionales, y de los procesamiento estadísticos de alta complejidad también ha sido parte de las nuevas propuestas. Desde una mirada sustentada en la Lingüística de Corpus y basada en el análisis multifactorial, Parodi (2005) clasifica los textos de acuerdo a su grado de densidad informativa. Esto con el propósito de desarrollar instrumentos de comprensión para evaluar el desempeño de lectores en ámbitos académicos. Para ello, y basado en la propuesta inicial de Biber (1988), somete los textos a un análisis que le permite identificar aquellos de baja densidad y aquellos de alta densidad informativa. Los textos de alta densidad informativa corresponden a aquellos que presentan mayor coocurrencia de ciertos rasgos lexicogramaticales, tales como nominalizaciones, frases preposicionales y cláusulas relativas con función de postmodificación, modo subjuntivo, etc. Por otra parte, los textos de baja densidad presentan la coocurrencia de otros rasgos lexicogramaticales, entre los que se encuentran pronombres en primera y segunda persona singular, tiempo pasado simple, verbos estativos, etc. Los resultados generales obtenidos demostraron que los lectores en ámbitos académicos tienen mayores problemas cuando se trata de enfrentar textos de alta densidad informativa.

Más recientemente, la lingüística computacional también ha sido incorporada a este tipo de investigaciones. Con una propuesta basada en un algoritmo, generado con base teórica en la semántica léxica, Louwerse y Jeuniaux (2009) pretenden determinar la similitud semántica entre los segmentos constituyentes de los textos y, de esta forma, identificar los textos más y menos cohesivos. Esta técnica es mejor conocida como Análisis Semántico Latente (Landauer & Dumais, 1997; Landauer, Foltz & Laham, 1998;) y consiste en un análisis algorítmico a partir de un corpus, el cual calcula la frecuencia de aparición de las palabras en determinados contextos.

Otro punto que ha sido considerado al momento de seleccionar los textos es su extensión. Diversos especialistas (Koda, 1987; Davey, 1988; Alderson, 2000) se han centrado en la relevancia que este aspecto puede tener a la hora de evaluar la

comprensión de textos escritos, dando origen a una gran controversia al respecto. Autores como Nuttall (1996), sostienen que a mayor extensión del texto, el nivel de logro disminuye. Con un mayor grado de precisión, Engineer (1977) propone que cuando los textos utilizados superan las 1.000 palabras, la medición se hace menos precisa. Desde nuestra perspectiva, el criterio de extensión no debe determinar la selección de los textos, sino que, muy por el contrario, creemos que la extensión debe estar supeditada al constructo que se desea medir. En términos más simples y a modo de ejemplo, cuando se desea medir relaciones semánticas implicadas en la generación de la coherencia local, el texto podría tener una extensión que permitiera establecer la o las relaciones que se pretende medir. Lo mismo puede ser aplicado cuando se desea medir constructos como el de Macroestructura, en cuyo caso la extensión del texto debería estar asociada a un cierre semántico.

Sin duda, las propuestas de caracterización de textos que hemos revisado han sido muy útiles para las investigaciones en que han sido utilizadas. No obstante ello, es preciso insistir en que cualquier caracterización pierde su valor si no es utilizada adecuadamente para medir los constructos teóricos que se desea medir y en virtud del propósito de la evaluación. En este punto, es preciso señalar que cualquier característica del texto, sea esta lingüística o estructural, debe ser entendida como la manifestación concreta del propósito comunicativo para el cual el texto fue creado. En este sentido, creemos que cuando se seleccionan los textos para un proceso de evaluación, estos no deben ser entendidos solo en cuanto a su dimensión lingüística, sino que, principalmente, como instancias de lengua y, por ende, como un acto de comunicación. Por esta razón, una selección de textos adecuada debería sustentarse en un estudio de corpus que permita dar cuenta no solo de las características lingüísticas de los textos, sino que también de sus rasgos contextuales, en tanto géneros discursivos.

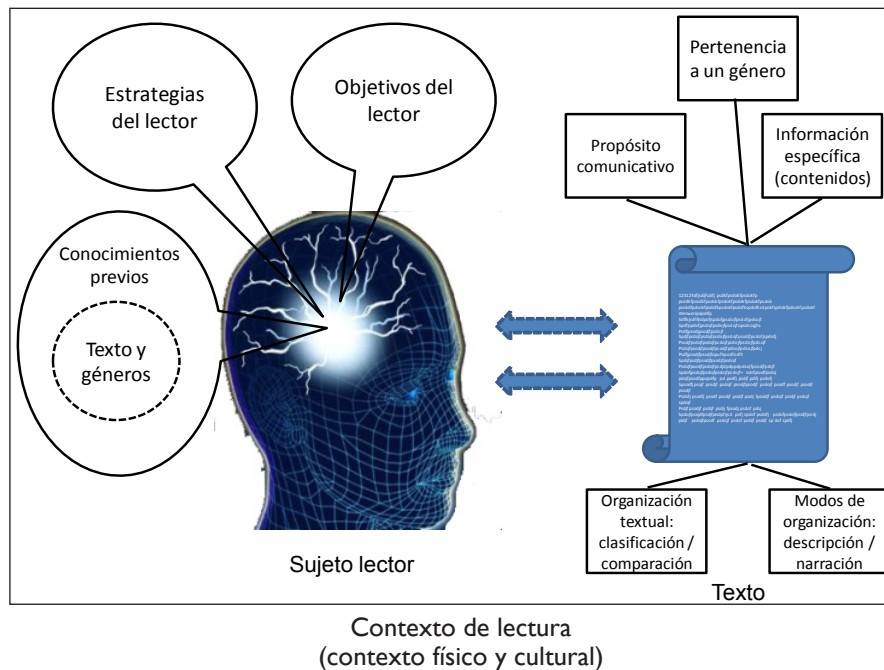
Desde nuestra mirada, cuando se trata de la selección de los textos, se debe tener en cuenta que la comprensión del discurso constituye un proceso situado e intencionado, que, por lo mismo implica una gran variación en las situaciones de lectura. De acuerdo a Parodi et al. (2010), cada vez que un individuo lee un texto lo hace porque necesita satisfacer ciertas necesidades que le impone el contexto. Tales necesidades determinarán el objetivo funcional del lector, entre los que se puede mencionar,

entretenerse, estudiar, deleitarse, etc. El objetivo funcional, a su vez, guiará los objetivos de lectura que el lector se imponga, entre los que se pueden mencionar, comprender la idea principal, identificar información específica, memorizar una definición, etc. Del mismo modo, el proceso estará guiado por dicho objetivo de lectura, el que determinará las estrategias a utilizar en la construcción del significado. Al mismo tiempo, es necesario considerar que los textos que se leen constituyen ejemplares de diferentes géneros discursivos, los que pueden ser entendidos como la abstracción convencionalizada de situaciones comunicativas particulares. Por ello, es posible sostener que existe una relación directa entre el género al cual pertenece el texto que se lee y los objetivos funcionales del lector (Magliano et al., 2007; Parodi et al., 2010). Esto implica que, generalmente, los objetivos de lectura, así como las estrategias implicadas estarán asociados a un género, su propósito comunicativo y sus características lingüístico-estructurales. Esto se grafica en la Figura 4.

Como se aprecia en la Figura 4, en cada proceso de lectura, la relación que se produce entre el lector y el texto implica también una relación entre sus conocimientos previos, sus objetivos funcionales, sus objetivos de lectura, sus estrategias y el género discursivo al cual el texto pertenece. Además, se debe tener en cuenta que los textos varían no solo en cuanto a sus temáticas, sino que también -como ejemplares de un género discursivo- varían en cuanto a sus propósitos comunicativos y, por ende, a sus características lingüísticas y estructurales (Parodi, Ibáñez & Venegas, 2010). Por esta razón, tanto la meta de lectura como las habilidades y estrategias necesarias para enfrentar textos pertenecientes a géneros diferentes varían. De este modo, cuando un estudiante de leyes debe prepararse para un examen, leerá el Código Civil con el objetivo de memorizar los artículos incluidos en tal examen y, por lo mismo, utilizará las estrategias correspondientes para lograr sus metas. Un cliente de un restaurant, a su vez, leerá el menú con el objetivo funcional de encontrar el plato que desea ordenar y el precio correspondiente. Por ello, su objetivo de lectura será identificar una información específica en el texto y no comprender la idea principal.

### 3.1.3. Una mirada integrada

De acuerdo a lo planteado hasta ahora, para que un instrumento dé cuenta de lo que se pretende medir, también es preciso establecer una relación



**Figura 4.** Proceso de lectura como evento social, lingüístico y cognitivo (Parodi et al., 2010).

coherente entre los elementos anteriormente descritos, es decir, resulta necesario que exista Coherencia Interna. Para ello, proponemos tres pasos básicos. En primer lugar, y en virtud del propósito de la evaluación, es necesario definir conceptualmente los constructos teóricos a medir. Luego, para la selección de los textos, así como para la definición operacional de los constructos teóricos, es necesario llevar a cabo un estudio de corpus que revele tanto cuáles son los géneros más característicos y adecuados para medir tales constructos, así como también, que permita conocer los rasgos léxico-gramaticales, semánticos y organizacionales de los mismos. La identificación y descripción de tales rasgos, hace posible construir un panorama de la configuración lexicogramatical particular, de las relaciones semánticas prototípicas y de la organización genérica. Por último, se deben seleccionar las técnicas de evaluación más adecuadas de acuerdo al constructo a medir, así como también a los rasgos de los géneros a los cuales los textos seleccionados pertenecen. Estos pasos permitirán no solo construir un instrumento que rescate las diferencias en términos de procesamiento que implica leer diferentes géneros, sino que también, operacionalizar los constructos teóricos de manera adecuada al contexto en que los géneros a utilizar circulan.

A modo de ejemplo, y con el propósito de presentar más claramente la variación anteriormente señalada, nos referiremos a dos géneros discursivos, a saber, el Cuento y el Manual. Estos ejemplos permiten observar no solo las diferencias genéricas, sino que también las diferencias en términos de procesamiento de la información que las diferencias genéricas imponen. Cualquier diseño y construcción de un instrumento debería cautelar tales diferencias para mantener su Coherencia Interna.

Los cuentos se encuentran entre los géneros que presentan una organización discursiva, predominantemente narrativa. Su estructura (Labov & Wallesky, 1967; van Dijk, 1980) está constituida por una presentación, un conflicto y un desenlace. De acuerdo a Zwaan y Radvansky (1998), estos están organizados en base a eventos, los cuales, pueden entenderse en términos de cinco dimensiones (espacial, temporal, causal, intencional y de los personajes). Generalmente, cuando los lectores se enfrentan a este género, lo hacen con el objetivo funcional de entretenerse, interiorizarse respecto de las costumbres de una época o sacar una moraleja. Por lo mismo, su objetivo de lectura consistirá en comprender la idea global, para lo cual será necesario utilizar diversas estrategias que le permitan comprender la secuencia general de los eventos, así como también, integrar la información en un todo coherente. De acuerdo a nuestra concepción

del proceso de comprensión, ello implica alcanzar no solo un nivel de representación del discurso equivalente a la Base Textual, sino que también al Modelo de Situación. Para alcanzar la Base Textual, es necesario que el lector establezca las relaciones semánticas necesarias que le permitan representar la coherencia local del texto. En el caso de los cuentos, y debido a la configuración del género, las relaciones semánticas subyacentes más frecuentes corresponderán a las de secuencialidad, causalidad, espacialidad, temporalidad y a correferencialidad (Parodi, 2011). Así también, la idea general de un cuento gira en torno a ciertos personajes, sus intenciones y las circunstancias en que los eventos tienen lugar. En este punto, el uso estratégico del conocimiento de mundo, así como de la estructura general del género juegan un rol fundamental en un proceso de comprensión profundo.

Por otra parte, un Manual, cuyo propósito comunicativo es eminentemente didáctico, presenta una organización discursiva predominantemente descriptiva. A diferencia de cualquier narración, los manuales no se organizan en eventos, sino que en base a tres movidas retóricas prototípicas, a saber, la definición de un concepto, la ejemplificación y la ejercitación (Parodi, 2010). Generalmente, cuando un lector enfrenta este género, lo hace con el objetivo funcional de estudiar o adquirir nuevos conocimientos para aplicarlos en situaciones reales. De ahí que, generalmente, su objetivo de lectura no consistirá, por ejemplo, en comprender la idea principal, sino que en entender con profundidad las definiciones de conceptos, así como también los procedimientos y procesos descritos. Si bien, en este caso, el lector no buscará comprender la idea principal del texto, comprender las definiciones y procedimientos también implica alcanzar un nivel de representación del discurso correspondiente al Modelo de Situación. Debido a la configuración del género, en el establecimiento de la coherencia local, las relaciones semánticas subyacentes más frecuentes que el lector deberá establecer corresponderán a las de secuencialidad, adversatividad, aditividad, y correferencialidad (Parodi, 2011).

Las diferencias que anteriormente presentamos respecto de los géneros discursivos, así como las implicancias que ellas tienen en un proceso de comprensión de textos escritos, demuestran que a la hora de seleccionar un texto para la construcción de un instrumento, es necesario cautelar que estos pertenezcan a géneros discursivos cuyas características permitan medir efectivamente

los constructos teóricos que se desea medir. Si bien constructos como Base Textual y Modelo de Situación parecen ser constructos que pueden ser evaluados a partir de textos pertenecientes a cualquier género, Macroestructura es un constructo que definitivamente puede ser mejor evaluado por medio de cuentos, noticias, novelas o fábulas, pero no por medio de diccionarios, manuales o recetas de cocina, por ejemplo. Así también, el constructo de Coherencia Local es un constructo que se podría entender como el más susceptible a la variación genérica. Esto debido a que la coherencia se establecerá a partir de relaciones semánticas de diferente tipo, las cuales, como anteriormente mostramos por medio de los ejemplos del Manual y del Cuento, se presentarán con mayor o menor frecuencia dependiendo del género. A continuación, en la Tabla 2, presentamos algunas posibles vinculaciones para establecer la Coherencia Interna.

A partir de las posibles combinaciones de elementos que proponemos en la Tabla 2, pretendemos mostrar con mayor explicitud las diversas posibilidades que un evaluador puede tener para establecer la Coherencia Interna de su instrumento. Esto, llegado el momento de decidir respecto de las técnicas de evaluación y los géneros a utilizar para medir, de manera más precisa, un constructo teórico determinado en virtud de un propósito de evaluación. Evidentemente, y según lo que planteáramos anteriormente, las posibilidades del evaluador variarán dependiendo tanto del propósito de la evaluación, como del contexto en el que esta se desarrolle. Un ejemplo claro de esto es lo que sucede cuando se llevan a cabo evaluaciones en ámbitos académicos. Allí, las disciplinas juegan un rol central. Esto pues, si bien en ocasiones, los propósitos de la evaluación pueden llegar a coincidir, los constructos teóricos a medir se operacionalizan de manera muy diferentes entre ciertas disciplinas y otras. Del mismo modo, los géneros discursivos prototípicos de una disciplina no lo serán necesariamente en otra.

Lo importante es poder cautelar y atender estas diferencias, para lo cual un trabajo de corpus previo es fundamental. Este tipo de indagación permite no solo identificar los géneros prototípicos en un contexto de lectura determinado, sino que también definir operacionalmente, de modo más preciso los constructos a medir. Asimismo, la selección de las técnicas a utilizar estará apoyada en datos empíricos, lo que resultará en una articulación más coherente y empíricamente sustentada.

**Tabla 2.** Posibilidades para establecer la Coherencia Interna.

| <b>Constructo teórico</b>    |  | <b>Técnica de evaluación</b>  | <b>Géneros más apropiados</b>  |
|------------------------------|--|---|--|
| <b>Definición conceptual</b> | <b>Definición operacional</b>  |   |  |
| Modelo de Situación          | <ul style="list-style-type: none"> <li>· Aplicar lo aprendido para resolver un problema</li> <li>· Aplicar lo aprendido para generar una analogía</li> </ul> | <ul style="list-style-type: none"> <li>· Pregunta abierta</li> <li>· Selección múltiple</li> <li>· Transferencia de registros semióticos</li> </ul> | <ul style="list-style-type: none"> <li>· Manual</li> <li>· Cuento</li> <li>· Receta de cocina</li> </ul>   |
| Macroproposición             | <ul style="list-style-type: none"> <li>· Expresar la idea principal</li> <li>· Proponer un título para el texto leído</li> </ul>                             | <ul style="list-style-type: none"> <li>· Pregunta abierta corta</li> <li>· Protocolo en voz alta</li> <li>· Verdadero y falso</li> </ul>            | <ul style="list-style-type: none"> <li>· Cuento</li> <li>· Noticia</li> <li>· Fábula</li> </ul>            |
| Macroestructura              | <ul style="list-style-type: none"> <li>· Producir un texto breve que dé cuenta de las principales ideas del texto leído</li> </ul>                           | <ul style="list-style-type: none"> <li>· Resumen</li> <li>· Resumen incompleto</li> </ul>   | <ul style="list-style-type: none"> <li>· Fábula</li> <li>· Artículo de investigación científica</li> </ul> |
| Coherencia local             | <ul style="list-style-type: none"> <li>· Establecer relaciones semánticas espaciales y temporales</li> </ul>   | <ul style="list-style-type: none"> <li>· Cloze</li> <li>· Pregunta abierta larga</li> </ul>   | <ul style="list-style-type: none"> <li>· Noticia</li> <li>· Cuento</li> </ul>                              |

Respecto de la integración que proponemos en la Tabla 2, resulta necesario señalar que esta rescata las posibles formas de establecer la Coherencia Interna. Por lo mismo, no incluye posibles propósitos de evaluación, ya que estos se encuentran en otra esfera, y que, como se grafica en la Figura 3, obedecen a necesidades contextuales de diferente índole, como podrían ser decidir respecto de políticas educativas, ubicar a los alumnos de acuerdo a sus capacidades, decidir respecto de las intervenciones pedagógicas más adecuadas en determinados contextos de aprendizaje, etc. (Ahumada, 1983; Alderson, 2000; Graesser, 2007; Magliano et al., 2007). No obstante lo anterior, es preciso insistir en la relevancia de la Coherencia Externa, ya que un propósito de evaluación corresponde al principio guía de un proceso evaluativo y, en ese sentido, determina las características de los instrumentos, así como la totalidad de los pasos involucrados en el procedimiento evaluativo, lo que incluye también la aplicación de los instrumentos y el análisis de los datos.

### Consideraciones finales

Como hemos venido sosteniendo, el éxito de un procedimiento evaluativo, es decir, que los resultados obtenidos en la evaluación permitan tomar decisiones adecuadas, depende, principalmente, de

que el instrumento utilizado mida con exactitud lo que se desea medir. Para ello, es necesario establecer una Coherencia Externa y una Interna. Teniendo esto en consideración, en este trabajo, hemos propuesto un marco para el diseño y construcción de instrumentos de evaluación de la comprensión de textos escritos, basado en nuestra concepción teórica del fenómeno. Hemos puesto especial énfasis en la vinculación de orden teórico-metodológico que se debe mantener entre los constructos teóricos a medir, las técnicas a utilizar y los textos a seleccionar para mantener la Coherencia Interna del instrumento. Para ello, hemos insistido en el valor que un estudio de corpus tiene para indagar respecto de diversos aspectos relevantes.

En esta propuesta no nos hemos referido a la asignación de puntaje, puesto que si bien entendemos que existe una jerarquía en cuanto a los procesos involucrados en la comprensión, también entendemos que el valor que el evaluador asignará a cada uno de ellos estará determinado por el propósito de la evaluación. Más precisamente, asumimos que en una evaluación que integre constructos teóricos asociados a determinados procesamiento psicolingüísticos, es necesario considerar la relevancia y dificultad que implica cada uno de ellos.

Como insistimos a lo largo del trabajo, los propósitos de evaluación constituyen los principios que guían el desarrollo de la totalidad del procedimiento evaluativo, desde su comienzo a su fin; esto es, desde el diseño de los instrumentos hasta la toma de decisiones que se pueden llegar a realizar en virtud del análisis de los datos obtenidos. Con nuestra propuesta, resumida en la articulación de elementos que ofrecemos en la Tabla 2, creemos que es posible mejorar el nivel de objetividad en las mediciones realizadas y, de este modo, alcanzar, de mejor manera el propósito de evaluación establecido por el especialista.

La revisión teórico-bibliográfica llevada a cabo para sustentar nuestra propuesta se focalizó principalmente en estudios relativos a la comprensión de textos escritos. De este modo, y si bien, abordamos ciertas temáticas relacionadas con una perspectiva, más bien psicométrica, es preciso señalar que nuestro interés ha sido aportar a la validez de los instrumentos desde el campo de la comprensión.

Creemos que, si bien esta propuesta puede ser mejorada y precisada, constituye un aporte para el

diseño y construcción de instrumentos de evaluación de la comprensión de textos escritos. Al mismo tiempo, asumimos que existen ciertos desafíos que debemos asumir como parte de una propuesta futura. Estos tienen que ver con profundizar en etapas posteriores al diseño y construcción y enfrentar lo que son los procedimientos involucrados en la validación. También, se hace necesario, para seguir avanzando en esta área, abordar en profundidad los procedimientos evaluativos de orden concurrente, en los cuales se integra mediciones de tipo *off line* con otras de naturaleza *on line*.

Por otra parte, actualmente, la variación disciplinar y los contextos de lectura, así como sus efectos en la comprensión del discurso escrito, se han transformado en foco de interés para quienes estudiamos el procesamiento del discurso escrito en contextos académicos (Parodi, 2005; Ibáñez, 2008; Santana, 2010). A partir de la presente propuesta, es posible enfrentar procesos de evaluación de la comprensión en ámbitos de especialidad, cautelando las diversas maneras en que las disciplinas determinan la manera en que se accede al conocimiento.

## REFERENCIAS BIBLIOGRÁFICAS

- Afflerbach, P. (2004). *National Reading Conference policy brief: High stakes testing and Reading achievement*. Oak Creek, WI: National Reading Conference.
- Ahumada, P. (1983). *Principios y procedimientos de evaluación educacional*. Valparaíso: EUVSA.
- Ahumada, P. (2002). *La evaluación en una concepción de aprendizaje significativo*. Valparaíso: EUVSA.
- Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.
- Alderson, J. C. & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, 35, 79-113.
- Anderson, N., Bachman, L., Perkins, K. & Cohen, A. (1991). An exploratory study into the construct validity of a reading comprehension test: Triangulation of data sources. *Language Testing*, 8(1), 41-66.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge, New York: Cambridge University Press.
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity (revised). En R. M. Paige (Ed.), *Education for the Intercultural Experience* (pp. 21-71). Yarmouth, Me: Intercultural Press.
- Biber, D. (1988). *Variation across speech and writing*. New York: Cambridge University Press.
- Bunge, M. (2004). *Emergencia y convergencia*. Barcelona: Gedisa.
- Carlisle, J. & Rice, M. (2004). Assessment of reading comprehension. En A. Stone, E. Silliman, B. Ehren & K. Apel (Eds.), *Handbook of language and literacy* (pp. 521-555). New York: Guilford Press.
- Chapelle, C. (1999). Validity in language assessment. *Annual Review of Applied Linguistics*, 19, (54-72).
- Cronbach, L. J. (1971). Test validation. En R. L. Thorndike (Ed.), *Educational Measurement* (pp. 443-507). Washington, D. C.: American Council on Education.
- Crossley, S. A., Allen, D. & McNamara, D. S. (2011). Text simplification and comprehensible input: A case for an intuitive approach, *Language Teaching Research* (Aceptada).
- Davey, B. (1988). Factors affecting the difficulty of reading comprehension items for successful and unsuccessful readers. *Journal of Experimental Education*, 56(1), 67-76.
- Diccionario de la Real Academia de la Lengua (DRAE) [en línea]. Disponible en: <http://buscon.rae.es/drae/>
- Eeds, M. (1988). Holistic assessment of coding ability. En S. M. Glazer, L. W. Searfoss & L. M. Gentile (Eds.), *Reexamining reading diagnosis: New trends and procedures* (pp. 48-66). Newark, DL: International Reading Association.
- Engineer, W. (1977). *Proficiency in reading English as a second language*. Tesis doctoral no publicada, University of Edinburgh.
- Fredrickson, J. W. (1984). The comprehensiveness of strategic decision processes: Extension, observations, future directions. *Academy of Management Journal*, 27, 445-466.

- Golden, R. M. & Goldman, S. R. (2006). An empirical feasibility study of the ARCADE system. En R. Sun (Ed.), *Proceedings of the 28th Annual Cognitive Science Society Conference* (pp. 1376-1381). Mahwah, NJ: Lawrence Erlbaum.
- Graesser, A. (2007). An introduction to strategic reading comprehension. En D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 3-26). Mahwah, NJ: Lawrence Erlbaum.
- Graesser, A., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371-395.
- Graesser, A. & Wiemer-Hasting, W. (1999). Situation models and concepts in story comprehension. En S. Goldman, A. Graesser & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence* (pp. 77-92). Mahwah, NJ: Lawrence Erlbaum.
- Graesser, A., Wiemer-Hastings, P. & Wiemer-Hastings, K. (2001). Constructing inferences and relations during text comprehension. En T. Sanders, J. Schilperoord & W. Spooren, (Eds.) *Text representation: Linguistic and psycholinguistic aspects* (pp. 249-271). Amsterdam: Benjamins.
- Hatch, E. & Lazaraton, N. (1991). *The research manual. Design and statistics for applied linguistics*. Massachusetts: Heinle & Heinle Publishers.
- Heaton, J. B. (1988). *Writing English Language Tests: A practical guide for teachers of English as a second or foreign language*. Londres: Longman.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Ibáñez, R. (2007). Cognición y comprensión. Una aproximación histórica y crítica al trabajo investigativo de Rolf Zwaan. *Revista Signos. Estudios de Lingüística*, 40(63), 81-100.
- Ibáñez, R. (2008). Comprensión de textos académicos escritos en inglés: Relación entre nivel de logro y variables involucradas. *Revista Signos. Estudios de Lingüística*, 41(67), 203-229.
- Kelly, E. L. (1927). *Interpretation of educational measurements*. Nueva York: Macmillan.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163-182.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Nueva York: Cambridge University Press.
- Kintsch, W. (2001). Predication. *Cognitive Science*, 25, 173-202.
- Kintsch, W. (2002). On the notions of theme and topic in psychological process models of text comprehension. En M. Louwerse & W. van Peer (Eds.), *Thematics Interdisciplinary Studies* (pp. 157-170). Amsterdam: Benjamins.
- Kintsch, W. & van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-341.
- Kintsch, W. & Rawson, K. A. (2005). Comprehension. En M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 209-226). Malden, MA: Blackwell.
- Klare, G. R. (1974-1975). Assessing readability. *Reading Research Quarterly*, 10(1), 62-102.
- Koda, K. (1987). Cognitive strategy transfer in second language reading. En J. Devine, P. Carrell & D. Eskey (Eds.), *Research in reading in English as a second language* (pp. 125-144). Washington, DC: TESOL.
- Labov, W. & Waletzky, J. (1967). Narrative analysis. En J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12-44). Seattle: University of Washington Press.

- Landauer, T. & Dumais, S. (1997). *A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge* [en línea]. Disponible en: <http://lsa.colorado.edu/papers/plato/plato.annote.html>
- Landauer, T., Foltz, P. & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25(2-3), 259-284.
- León, J.A., Otero, J. & Graesser, A. C. (Eds.) (2002). *The psychology of science text comprehension*. Mahwah, NJ: Lawrence Erlbaum.
- Leslie, L. & Caldwell, J. (2009). Informal and formal measures of reading comprehension. En S.E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 403-427). Nueva York: Routledge.
- Lomicka, L. (1998). To gloss or not to gloss: An investigation of reading comprehension online. *Language Learning and Technology*, 1(2), 41-50.
- Louwerse, M. (2002). An analytic and cognitive parameterization of coherence relations. *Cognitive Linguistics*, 12, 291-315.
- Louwerse, M. (2004). Un modelo conciso de cohesión en el texto y coherencia en la comprensión. *Revista Signos. Estudios de Lingüística*, 37(56), 41-58.
- Louwerse, M. & Van Peer, W. (2002). *Thematics: Interdisciplinary studies*. Amsterdam: Benjamins.
- Louwerse, M. & Graesser, A. (2006). Macroestructura. En E. K. Brown (Ed.), *Encyclopedia of language and linguistics* (pp. 426-429). Amsterdam: Elsevier.
- Louwerse, M. & Jeuniaux, P. (2009). Computational psycholinguistic techniques to measure cohesion in discourse. En J. Renkema (Ed.), *Discourse of course* (pp. 213-226). Amsterdam: Benjamins.
- Magliano, J. P. & Graesser, A. (1991). A three-pronged method for studying inference generation in literary text. *Poetics*, 20, 193-232.
- Magliano, J. P. & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure. *Cognition and Instruction*, 21, 251-283.
- Magliano, J. P., Millis, K. K., Ozuru, Y. & McNamara, D. (2007). A multidimensional framework to evaluate reading assessment tools. En D. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 107-136). Mahwah, NJ: Lawrence Erlbaum.
- Mannes, S. & St. George, M. (1996). Effects of prior knowledge on text comprehension: A simple modeling approach. En B. K. Britton & A. C. Graesser (Eds.), *Models of Understanding Text* (pp. 115-139). Mahwah, NJ: Lawrence Erlbaum.
- McKoon, G. & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99, 440-446.
- McNamara, D. (2004). Aprender del texto: Efectos de la estructura textual y las estrategias del lector. *Revista Signos. Estudios de Lingüística*, 37(55), 1-12.
- McNamara, D. & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, 22, 247-288.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18(2), 5-11.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.
- Nuttall, C. (1996). *Teaching reading skills in a foreign language*. Londres. Heinemann.

- O'Brien, E. J., Rizzella, M. L., Albrecht, J. & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24, 1200-1210.
- O'Brien, E., Cook, A. & Guéraud, S. (2010). Accessibility of outdated information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(4), 979-991.
- O'Reilly, T., Sabatini, J., Bruce, K. & Sukkarieh, J. (2011). What multiple choice tests of reading ability don't tell you: Evidence of reflection and knowledge integration in high quality constructed responses. Educational Testing Service, Princeton, NJ, U.S.A.
- Paris, S. G. & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36, 89-101.
- Parodi, G. (1991). Comprensión inferencial y estructura textual. *Revista Signos. Estudios de Lingüística*, 24(30), 81-89.
- Parodi, G. (2000). La conexión lectura-escritura: Estudio discursivo cognitivo. *Revista de Lingüística y Literatura*, 33, 109-127.
- Parodi, G. (2002). Comprensión y producción lingüísticas: Una nueva mirada al procesamiento del discurso escrito. *Revista Versión*, 11, 59-97.
- Parodi, G. (2005). La comprensión del discurso especializado escrito en ámbitos técnico profesionales: ¿Aprendiendo a partir del texto? *Revista Signos. Estudios de Lingüística*, 38(58), 221-267.
- Parodi, G. (2007). El discurso especializado escrito en el ámbito universitario y profesional: Constitución de un corpus. *Revista Signos. Estudios de Lingüística*, 40(63), 147-178.
- Parodi, G. (2010). La organización retórica del género manual a través de cuatro disciplinas: ¿Cómo se comunica y difunde la ciencia en diferentes contextos universitarios? *Revista Boletín de Lingüística*, 22(33), 43-69.
- Parodi, G. (2011). La Teoría de la Comunicabilidad: Notas para una concepción integral de la comprensión de textos escritos. *Revista Signos. Estudios de Lingüística*, 44(76), 145-167.
- Parodi, G., Ibáñez, R. & Venegas, R. (2010). Discourse Genres in PUCV-2006 Corpus of Academic and Professional Spanish. Criteria, Definitions and Examples. En G. Parodi (Ed.), *Discourse genres in Spanish: Academic and professional connections* (pp. 37-63). Amsterdam: Benjamins.
- Parodi, G., Peronard, M. & Ibáñez, R. (2010). *Saber leer*. Madrid: Aguilar.
- Pérez Zorrilla, M. J. (2005). Evaluación de la comprensión lectora: Dificultades y limitaciones. *Revista de Educación*, 121-138.
- Perfetti, C. A., Yang, C-L. & Schmalhofer, F. (2008). Comprehension skill and word-to-text processes. *Applied Cognitive Psychology*, 22(3), 303-318.
- Peronard, M. (1989). Estrato social y estrategias de comprensión de lectura. *Lenguas Modernas*, 16, 19-32.
- Peronard, M. (1992). La comprensión de textos escritos como proceso estratégico. En A. Bocaz (Ed.), *Actas del primer simposio sobre cognición, lenguaje y cultura: Diálogo transdisciplinario en ciencia cognitiva* (pp. 89-102). Santiago: Editorial Universitaria.
- Peronard, M. (1999). Metacognición y Conciencia. En G. Parodi (Ed.), *Discurso, Cognición y Educación* (pp. 43-57). Valparaíso: EUVSA.
- Peronard, M. (2007). Lectura en papel y en pantalla de computador. *Revista Signos. Estudios de Lingüística*, 40(63), 179-195.

- Peronard, M. & Gómez Macker, L. (1985). Reflexiones acerca de la comprensión lingüística: Hacia un modelo. *Revista de Lingüística Teórica y Aplicada (RLA)*, 23, 19-32.
- Peronard, M., Gómez Macker, L., Parodi, G., Núñez, P. & González, J. (1997). *Programa L y C: Leer y Comprender*. Santiago de Chile: Andrés Bello.
- Peronard, M., Crespo, N. & Velásquez, M. (2000). *La evaluación del conocimiento metacomprensivo en alumnos de educación básica*. *Revista Signos. Estudios de Lingüística*, 33(47), 167-180.
- Peronard, M., Crespo, N. & Guerrero, I. (2001). La metacomprensión en escolares chilenos de educación básica. *Revista Signos, Estudios de Lingüística* 34(49-50), 149-164.
- Peronard, M., Crespo, N. & Velásquez, M. (2002). Conocimiento metacognitivo del lenguaje escrito: Instrumento de medición y fundamentación teórica. *Infancia y aprendizaje*, 25(2), 131-145.
- Peronard, M., Velásquez, M., Ibáñez, R., Órdenes, J. & Alonzo, T. (2006). *Guiones metodológicos para desarrollar estrategias de comprensión y producción de textos escritos*. Valparaíso: EUVSA.
- Rauch, D. & Hartig, J. (2010). Multiple-choice versus open-ended response formats of reading test items: A two-dimensional IRT analysis. *Psychological Test and Assessment Modeling*, 52(4), 354-379.
- Rupp, A., Ferne, T. & Choi, H. (2006). How assessing reading comprehension with multiple-choice questions shapes the construct: a cognitive processing perspective. *Language Testing*, 23, 441-474.
- Sanders, T. & Spooren, W. (2001). Modeling causal and contrastive connectives: On domains, subjectivity and mental spaces. *International Cognitive Linguistics Conference*. Santa Barbara, CA: USA.
- Sanders, T. & Pander, H. (2006). Cohesion and Coherence: Linguistic approaches. En K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (pp. 591-595). Londres: Elsevier.
- Santana, A. (2010). *Nivel de dominio del inglés, multimodalidad y nivel de comprensión en segunda lengua de textos académicos escritos en inglés: Una aproximación al fenómeno en el ámbito de la Física*. Tesis para optar al grado de Magíster en Lingüística Aplicada, Pontificia Universidad Católica de Valparaíso, Chile.
- Shepard, L. (1993). Evaluating test validity. *Review of Research in Education*, 19, 405-50.
- Simon, H. (1957). *A Behavioral Model of Rational Choice*, in *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. Nueva York: Wiley.
- Thorndike, E. (1904). *An introduction to the theory of mental and social measurements*. Nueva York: Teachers College, Columbia University.
- Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. Chicago: The University of Chicago Press.
- van Dijk, T.A. (1980). *Macrostructures*. Hillsdale, NJ: Lawrence Erlbaum.
- van Dijk, T. & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. Nueva York: Academic Press.
- Walter, C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Psycholinguistics*, 25, 315-339.
- Weir, C. J. (1990). *Communicative language testing*. Londres: Prentice Hall.
- Weir, C. J. (2005). Quantifying Test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, 19(1), 231-240.

- Zwaan, R. A. & Madden, C. J. (2004). Updating situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 283-288.
- Zwaan, R. A. & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162-185.

\* Investigación financiada por los Proyectos FONDECYT 11080265 y 1090030.