

# *Estudios mixtos: la función ancilar del enfoque cuantitativo en la investigación educativa*

*Dalia Reyes Valdés<sup>1</sup>*

*José R. Reyes Valdés<sup>2</sup>*

*María Eugenia Flores Treviño<sup>3</sup>*

## **Introducción**

La relevancia de los enfoques metodológicos mixtos es, además de vigente, premonitoria desde que los estudios sociales pasaron de ser un oficio a conformar procesos científicos para la toma de decisiones en lo micro y macrosocial. El objetivo de esta propuesta es detallar un caso en donde los estudios cuantitativos se convirtieron no en el desenlace de una investigación, sino en el punto de partida de una novedosa metodología para capitalizar los corpus lingüísticos de estudiantes de secundaria al correlacionarlos con datos censales en México.

La metodología propuesta se basó en la minería de datos, con resultados de alto valor para la adecuación de estrategias en el aula, la atención a casos de alumnos en riesgo y la propuesta de políticas públicas. La viabilidad de esta iniciativa correlacional es una de las conclusiones relevantes del avance de tesis *La escuela secundaria como reguladora de los factores discursivos correlativos entre disponibilidad léxica y movilidad social* en la Facultad de Filosofía y Letras de la Universidad Autónoma de Coahuila en trabajo colaborativo con la Escuela Normal Superior del Estado de Coahuila, ambas en México.

---

<sup>1</sup>Dalia Reyes Valdés. Escuela Normal Superior del Estado de Coahuila-Universidad Autónoma de Nuevo León, México. Doctorante en la FFyL UANL acentuación en estudios de la educación, especialista en análisis del discurso docente. [dalia68reyes21@gmail.com](mailto:dalia68reyes21@gmail.com)

<sup>2</sup>José R. Reyes Valdés. Centro de Investigación en Matemática Aplicada, Universidad Autónoma de Coahuila, México. Doctor en Economía Regional, investigador especialista en ciencia de datos. [jose.reyes.valdes@uadec.edu.mx](mailto:jose.reyes.valdes@uadec.edu.mx)

<sup>3</sup>María Eugenia Flores Treviño. Facultad de Filosofía y Letras UANL, México. Doctora en Humanidades y Artes se especializa en análisis del discurso. [maria.florestre@uanl.edu.mx](mailto:maria.florestre@uanl.edu.mx)

## **De los enfoques mixtos a los métodos colaborativos**

La discusión académica de un estudiante con su asesor para dejar claramente establecido en enfoque metodológico al cual apegarán una tesis doctoral es un capítulo determinante en las reuniones tutoriales: no solo determinará los muros fronterizos para el doctorante sino las riberas entre las cuales deberá acotarse el avance del estudio. Ahora bien, si los diversos enfoques se han desarrollado para conocer la realidad y tipificarla como un problema por estudiar, cabría cuestionarse si no debería ser esta, y no la suposición teórica, quien determine los derroteros de adquisición de datos y sus procesos para usar, descartar, amalgamar o inaugurar nuevas metodologías de investigación.

La discusión por la supremacía de los enfoques metodológicos está sobreseída, puesto que la investigación ha superado ya el prejuicio sobre la subvaloración del ejercicio fáctico en los estudios de ciencias sociales. Se ha dejado de llamar dato blando al que construye el edificio teórico para la toma de decisiones sobre los derroteros sociales, tanto como se canceló la certeza de creer única a la vía cuantitativa para conocer una realidad certera, que ni es del todo medible ni replicable: cada vez más se encuentran definitivas las realidades individualizadas o socializadas, principalmente en la investigación educativa.

En el proceso de investigación educativo que detonó este documento se mostró la urgencia por un trabajo interdisciplinar sin establecer parcelas propias ni capítulos en propiedad; al mismo tiempo requirió el replanteamiento de una reflexión metodológica para el establecimiento de un método colaborativo emergente que dio resultados positivos tanto para la política pública como para la toma de decisiones en el aula.

Los métodos establecen pasos y reglas para hacer una manipulación inteligente de la realidad categorizada como problema, afirman Aguilera y Rina (2013); sin embargo, esta forma de proceder responde a una internalización reflexiva del conocimiento y sus valores teleológicos. De lo anterior resulta el inalienable involucramiento de la postura ontológica y epistemológica del investigador quien, en efecto, parte de lineamientos establecidos que son viables der ajustados a la realidad que estudia.

El método para la construcción de corpus con índices de disponibilidad léxica está claramente establecido hace más de 30 años; por otro lado, las corrientes para el análisis discursivo han sido

caracterizadas y delimitadas para estudios específicos con finalidades diversas. La emergencia metodológica se presentó al encontrar que ambos marcan como final del proceso un paso que, en realidad, es inicio para el siguiente y la propuesta de investigación correlacional debió zanjar ese espacio.

### **La estadística: el reto y la posibilidad**

Como requisito para obtener el título de doctor, algunos programas exigen realizar estancias de investigación. El sentido común predomina en los estudiantes: se busca tomar cursos del nivel con estrecha relación temática o bien, participar en proyectos colaborativos con especialistas de su misma disciplina y estudio; lo atípico es, como en el caso de este trayecto de investigación, intentar ser aceptado por un sociólogo estadista para hablar sobre la construcción del lenguaje, sus causas y sus consecuencias.

El trayecto doctoral para la tesis *La escuela secundaria como reguladora de los factores discursivos correlativos entre disponibilidad léxica y prospectiva de movilidad social* se inició con la idea fija -sin muchos fundamentos metodológicos- de que un mismo estudio podía encontrar el nodo común entre los propósitos educativos de un nivel, el ejercicio educativo en la escuela secundaria, la formación del lenguaje en ese trayecto y el impacto social presente y futuro que esta construcción podría tener.

Reconocer las ligas que unen y extienden los temas sociales, y estar abierto a las posibilidades correlacionales de dos métodos con diferente cuna, tendría que ser un requisito para el perfil de ingreso en el campo investigativo. En realidad, hace tiempo que se llevan a cabo, pero, al parecer, hay prejuicios todavía latentes para catalogarlos como relevantes; el mundo de los lingüistas no parece ser diferente.

En 1996, López-Morales atisba las emergencias que, *a posteriori*, requerirían la intervención de especialistas en áreas disciplinares diferentes a aquellos dedicados al estudio de la lengua. En su artículo *Los estudios de disponibilidad léxica: pasado y presente* acepta las falencias de su ejercicio en ciernes por sistematizar los estudios:

Charles Müller, sin embargo, había advertido de la conveniencia de tomar en consideración el orden de aparición de las palabras en las listas, ejercicio

que daría una versión más refinada de la realidad: varias palabras podían alcanzar la misma frecuencia y no tener el mismo índice de disponibilidad (1996, p. 246).

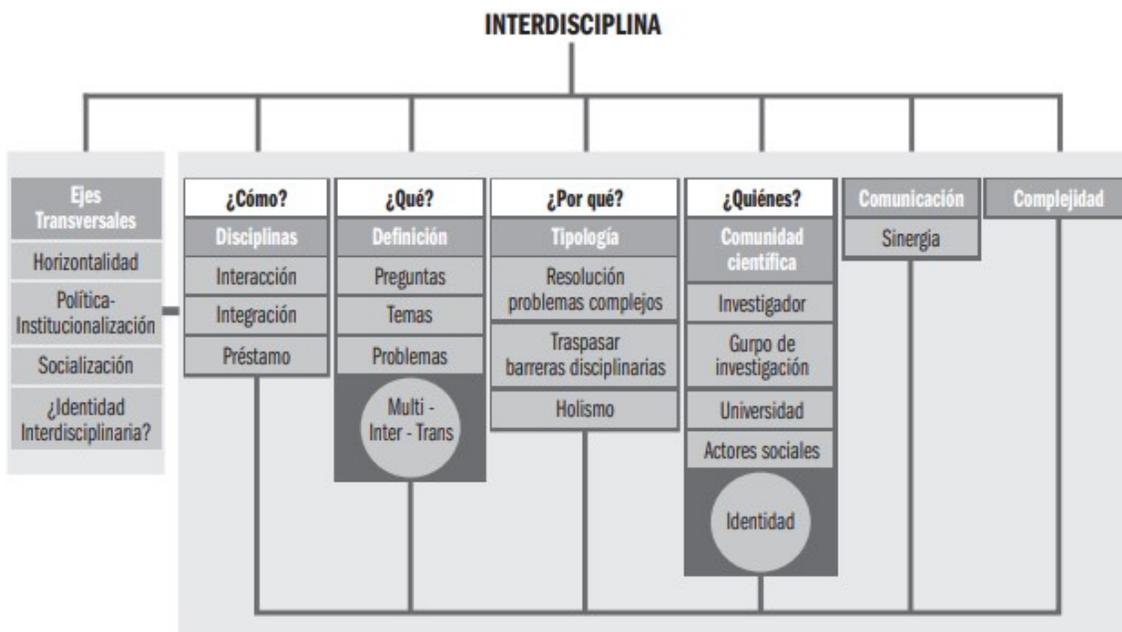
Al mismo tiempo que se extiende en la perspectiva futura, cierra sus cotos al ámbito cuantitativo con la afirmación de que el índice de disponibilidad léxica es “una medida que une criterios de frecuencia y orden, final matemático de un camino axiomático muy riguroso” (López-Morales, 1996, p. 248). De todos modos, el plano cuantitativo resultó no ser ni el fin ni la finalidad, en realidad estaban abriendo posibilidades rizomáticas para diversas disciplinas con objetivos múltiples.

El avance de tesis que genera este texto, por dar un ejemplo de la diversificación de finalidades, enlaza los datos numéricos de disponibilidad léxica en un estudio longitudinal de informantes durante su trayecto en la escuela secundaria, y se validan los vocablos atípicos como generadores cualitativos para hacer una correlación estadística contextual con las zonas geográficas de la muestra cerrando el estudio con el análisis discursivo de entrevistas en consecuencia. Matemáticas las hay -tanto como Sociología- no como un punto de llegada, sino como un puente que coloca a la disciplina en una muy importante función ancilar, desde sus sub disciplinas y como guía para la iteración sobre el tejido que forma un corpus. En todo caso, este documento es un *continuum* del publicado hace más de 20 años por López-Morales: aquí se trata sobre el futuro de los estudios de disponibilidad léxica.

### **El discurso: una posibilidad para la estadística**

Hoy lo que se ha puesto sobre la mesa es la necesidad emergente de diseñar métodos pertinentes para trabajar en el ambiente interdisciplinar. Vienni (2015) hace un interesante extracto de conceptos viables de convertirse en las bisagras que abran la puerta al concubinato indiscriminado, pero sistemático, de diversas disciplinas (*Cuadro 1*).

**Cuadro 1. Conceptos relevados en la literatura científica para la interdisciplina.**



Fuente: Vienni, 2015, p. 45

Los cierres de columna: sinergia, identidad, multi-inter-trans y préstamo parecen las palabras clave para el ejercicio de un trabajo colaborativo. Al mismo tiempo refieren a las acciones por realizar entre los participantes como las dimensiones disciplinares que deberán conservarse en el proceso, pues una producción entre varias disciplinas no implica la despersonalización –ni devaluación- de ninguna.

El documento de López-Morales (1996) es muy valioso como parteaguas en el reconocimiento de las disciplinas que empezarían a involucrarse en el cálculo, manejo y aplicación de los corpus de disponibilidad léxica. La Universidad Autónoma de Zacatecas (Hernández 2013; López, 2003; Cortez 2016,) ha hecho un trabajo descomunal con la implantación de un estudio sistematizado en sus instrumentos, el pulimento de las fórmulas matemáticas –todavía bajo la lupa de la perfección-, la creación de un programa específico para el cálculo, las aplicaciones, las correlaciones y, sobre todo, el aporte desinteresado de resultados y herramientas a los estudiosos de la lengua. Es importante mencionar que los índices de disponibilidad léxica se han

utilizado, prioritariamente, para la generalización, pero en este estudio la particularidad es relevante para la toma de decisiones en el aula.

El análisis de sentimiento, una vertiente de la lingüística computacional, procesados con software libre de estadística es una tendencia mundial para quienes se inclinan por la técnica de minería de datos, es decir la búsqueda de lo no evidente desde los datos visibles. La propuesta metodológica planteada en la tesis doctoral aludida está directamente relacionada con ambas: los corpus de índices de disponibilidad léxica son la veta que anuncia tiros significativos de fondo no identificado; ahora bien, en lo que refiere a análisis de sentimiento, las metodologías de análisis discursivo tienen bastante que aportar a los estadísticos abocados en este momento a desvelar misterios en las palabras de millones de tweets (Baviera, 2017), por dar un ejemplo.

La coyuntura en el lenguaje se busca desde las condiciones de producción. Dice Courtine (1981) que se trata de una mediación ideológica, con su periodo de caos creativo, cuyos orígenes están en la psicología social, en la sociolingüística y en los implícitos del mismo discurso, es decir, en la materia oscura viable de comprenderse a partir de la identificación de índices discursivos, actos de habla (Searle, 2015) y puesta en marcha de las funciones del lenguaje (Jakobson, 1984). En el caso que aquí ocupa, los vocablos obtenidos en los corpus fueron el denominador de un discurso por conocer que se amplió, posteriormente, mediante entrevistas semiestructuradas; si se considera una muestra de 116 informantes, hubiese sido una tarea ardua -que en el pasado se realizaba- acudir de forma directa a las entrevistas para luego seleccionar informantes con marcas de prospectiva de movilidad social sin mediar el cálculo de disponibilidad.

Al partir del corpus indicial se identificaron los casos con valores extremos (cuantitativo) y los casos atípicos (cualitativo). La coyuntura entre los vocablos arrojados por los informantes se buscó, en el análisis discursivo, desde sus implicaciones contextuales -cultural, familiar, religioso, económico- y cotextuales -los discursos periféricos que lo alimentan. De esta manera, el enfoque cuantitativo arrojó datos interpretables cuyo análisis desglosa categorías cuantificables que, a su vez, abrieron nuevas incógnitas cuantitativas, formando un bucle interminable y dinámico.

La ciencia de datos -que contiene a la minería como técnica- además de buscar la información no evidente también pretende utilizarla para predecir comportamientos; si bien se inclina más

al servicio de los ámbitos económicos o políticos, ese interés se manifiesta también en los estudios de léxico, mejor planteado por López (2003) cuando habla de una forma para mejorar la posibilidad es salir de una asociación específica inter corpus para ir a una serie de asociaciones que permitan explicar los factores que las determinan.

## **Defensa bilateral**

La urgencia por la interdisciplinariedad muestra la inexistente contraposición entre lo cuantitativo y lo cualitativo. El malentendido referente a la falsa carencia de categorías en los estudios sociales o la imposibilidad de aplicar en ellos teorías abstractas quedó atrás al comprender que las metodologías para estos no están basadas en creencias.

Lo cualitativo se categoriza; al hacerlo, se cuantifica a partir de la codificación, lo que abre paso a la utilización de esos códigos como base de modelos estadístico-matemáticos para luego regresar a retroalimentar el contexto de aplicación. En realidad, el nacimiento de una investigación cuantitativa en el ámbito educativo surge de lo cualitativo, pero categorizado que alcanza la posibilidad de procesamientos diversos y enriquecidos de los mismos datos. Todo es un dato y, realidad, los datos duros no lo son tanto: cualquiera que se genere es una transformación numérica de algo que se intenta medir y el resultado cambiará si se modifica el lugar, la hora, el contexto, el estado emocional.

El término *boot strap* es cercano a una tendencia interdisciplinar como la que aquí se defiende: sacar la mayor información de pocos datos, extraerla de donde aparentemente no había más. De la mano, un estudiante de Ciencias Sociales asesorado por uno de los autores de este texto, busca aplicar modelos de la física que se aplican al comportamiento socioeconómico: entropía, teoría del caos, modelos gravitatorios que tienen que ver tanto con la gravitación de las estrellas como con economía.

Los estudios cuantitativos revelan patrones que pueden estar ocultos con nuestra limitación, se ven correlaciones que nuestro cerebro prejuiciado puede ver por el contexto personal y entonces, al enlazarse con procesos analíticos interpretativos se descubren asociaciones no previstas. El antecedente más cercano es el análisis de sentimiento utilizado para estudiar comportamientos

en redes sociales, y el más reciente, este que aquí se describe: encontrar comportamientos atípicos a partir de valores léxicos numéricos e inferir acciones *a posteriori*.

## **Bibliografía**

- Aguilera, H. & Rina. M. (2013) Identidad y diferenciación entre método y metodología. *Estudios Políticos*, 28 (No. 47), 81-103.
- Baviera, T. (2017). Técnicas para el Análisis de Sentimiento en Twitter: Aprendizaje Automático Supervisado y SentiStrength. *Revista Dígitos. 1* (No. 1), 33-50.
- Cortez, G. (2016). *Una aplicación de la disponibilidad léxica: los autores de obras literarias más disponibles entre los estudiantes de diez licenciaturas de la UAZ*. Zacatecas, México: UAZ.
- Courtine, J. (1981). Análisis del discurso político. (El discurso comunista dirigido a los cristianos). *Revista Langages*, 2 (N° 62).
- Hernández, M. (2013). Léxico disponible de niñas y de niños de 6° de primaria de la ciudad de Zacatecas. En López, J. (Ed.), *Recuento y perspectiva de la enseñanza del español* (pp. 106-113). México: Edere.
- Jakobson, R. (1984). *Lingüística y poética. Ensayos de lingüística general*. Barcelona: Ariel
- López-Morales, H. (1996) Los estudios de disponibilidad léxica, pasado y presente. *Boletín de Filología de la Universidad de Chile Homenaje a Rodolfo Oroz* 35 (No. 1), 245-259.
- López, J. (2003). *Qué te viene a la memoria*. Zacatecas, México: UAZ
- Searle, J. (2015). *Actos de habla*. Madrid: Cátedra.
- Vienni, B. (2015). Los estudios sobre interdisciplina: construcción de un ámbito en el campo de ciencia, tecnología y sociedad. *Revista Redes*, 21 (No.41), 141-175.