Del punto al área: consideraciones desde la teoría de la complejidad para representar el aprendizaje léxico

Rodrigo Ibarra González
Universidad Autónoma de Querétaro
Ignacio Rodríguez Sánchez
Universidad Autónoma de Querétaro
Antonio Rico Sulayes
Universidad de las Américas Puebla

1. La ciencia como sesgo de realidad

Aunque la percepción más común de la ciencia es la de una actividad homogénea, existen otras tradiciones y prácticas muy distintas en torno a tal concepción. Habermas (2002) afirma que los métodos para hacer ciencia y los fines que ésta persigue se hallan estrechamente vinculados y que, por tanto, el método cambia según la finalidad. Tanto la ciencia como la tecnología son constructos culturales respaldados por una comunidad. Esta validación en buena medida proviene de la repetición de prácticas y acciones comunes, las cuales legitiman las formas de hacer ciencia.

Una interpretación de cómo hacer ciencia proviene del método científico. El éxito de este método se explica porque mecaniza la realidad y genera predicciones a partir de afirmaciones del tipo "Si X sucede en condiciones A, entonces Y sucederá en condiciones B". Estas oraciones condicionales, o premisas seguidas de conclusiones, las cuales son típicas de la lógica clásica, se transforman en el ámbito científico en causalidades, en leyes que resumen un conocimiento. Lograr esta mecanización de reglas de cierto conocimiento requiere que la realidad y las relaciones que en ellas se dan se simplifiquen en la mayor medida posible. A aproximadamente 400 años del uso recurrente del método científico (García, 2006), conviene pensar en las limitantes de un método que sistemáticamente disgrega la realidad. ¿Qué perdemos en aras del mecanicismo recién ejemplificado?

Feyerabend (1986) sugiere que un medio complejo requiere de explicaciones y métodos complejos. Un medio complejo desafía los análisis basados en procedimientos y metodologías preestablecidas que no consideran condiciones cambiantes en la historia. Al repetir constructos teóricos hegemónicos sin un proceso de cuestionamiento, los hechos adversos a la narrativa dominante tienden a recibir explicaciones un tanto burdas. En caso de teorías, continúa el autor, la unanimidad de opiniones no cuestionadas ni cuestionables es un hecho tan peligroso como la visión religiosa.

En este artículo ilustraremos cómo, desde la teoría de sistemas dinámicos, la recopilación compleja de datos permite realizar un análisis sofisticado del proceso de aprendizaje de vocabulario de segundas lenguas. Ello aporta una visión de la realidad más compleja, rica y con mayor potencial explicativo. Para esto, el artículo se desarrolla en tres etapas.

En primer lugar, presentamos una revisión de la teoría de sistemas dinámicos y sus aplicaciones en el ámbito del aprendizaje, específicamente en el aprendizaje de vocabulario de segundas lenguas.

En segundo lugar, explicamos cómo la forma en que se recogen, seleccionan y leen los datos concede hacer lecturas más complejas del conocimiento. Aquí haremos hincapié en distintas maneras de visualizar la información que influye en cómo entendemos el conocimiento. En este caso concreto, exponemos las limitaciones y algunas propuestas de representaciones gráficas, pues éstas resultan ilustrativas de algunos sesgos, restricciones

y fortalezas del método científico. Por ejemplo, argüiremos que la gráfica es una abstracción de un segmento de la realidad de interacciones entre un cúmulo de variables en una situación determinada. Entonces, analizamos diferentes tipos de gráficas, de menor a mayor complejidad, que suelen usarse para hacer representaciones; para ello recurriremos al ejemplo del desarrollo léxico de un individuo.

En tercer lugar, discutimos algunas de las implicaciones que se desprenden de abordar la recolección, el procesamiento y la interpretación de datos complejos.

2. La teoría de sistemas dinámicos (TSD) y la adquisición y el desarrollo de la lengua

Durante mucho tiempo el aprendizaje se ha concebido (igual que la historia y el conocimiento) como una actividad lineal, continua e incremental. Utilizando la manida metáfora de la concepción bancaria o alimentaria del conocimiento (Freire, 1970), el aprendizaje de la lengua se conceptualiza como una alcancía. Cada concepto o información transmitida es como una moneda que se agrega o "alimenta" la alcancía. El aprendizaje adquirido entonces es el resultado de la suma de todos los conceptos. No hay pérdidas ni saltos. Posiblemente una de las señales más claras de esta visión en la que el aprendizaje de la lengua se da como la suma de diferentes valores sea la que proviene del concepto zona de desarrollo próximo (ZPD, por sus siglas en inglés) creado por Vygotsky (1978). El autor define esta noción como

La distancia entre el nivel de desarrollo actual determinado por sus capacidades para solucionar problemas y el nivel potencial de desarrollo, determinado a través de las capacidades de solucionar problemas bajo la guía de un adulto o la colaboración con pares más capaces.

Según esta noción, se espera que el aprendiz reciba un *input* un poco más complejo del que por lo común maneja, para que

estas instancias de lengua, en el caso del aprendizaje lingüístico, sirvan de escalones en su crecimiento. La forma en la que se resume dicha idea es con una fórmula de exactitud casi matemática: I+1. Tanto la idea de la ZPD como del I+1 repercuten en el aprendizaje incremental tipo alcancía. Si el proceso de adquisición inicia en un punto X, al ser expuesto al lenguaje tipo I+1, se hallará en un periodo de adquisición X+1, seguido de X+2, X+3 y así consecutivamente. El motivo por el cual los procesos de aprendizaje se conceptualizan como continuamente incrementales se debe a que implícitamente se cataloga al lenguaje como un sistema simple.

Los sistemas simples son, en esencia, sistemas cuyos componentes tienen una relación parcial o total con el sistema y cierto grado de independencia (García, 2006). En el caso de la lengua, elementos tales como los morfemas son identificables y tienen un significado altamente estable, lo cual asegura su independencia y relación con el sistema lingüístico. Debido a que los componentes de los sistemas simples son, en principio, independientes, es posible establecer relaciones causales. El enfoque de la lengua como un sistema simple es consistente con los fines del método científico. Este punto de vista y método hunde sus raíces en el positivismo, e intenta explicar y predecir lo que pasa en el mundo social al buscar regularidades y relaciones causales entre sus elementos constituyentes. Este proceso explicativo es claramente reduccionista, pues el todo es segmentado continuamente en sus constituyentes (Hirschheim, 2002).

Además de esta conceptualización del lenguaje como sistema simple, hay quienes consideran que el lenguaje posiblemente tenga otra naturaleza. Algunos sugieren una postura alternativa en la que las palabras no son consideradas objetos —componentes en un listado mental— sino procesos, como lo sugiere un acercamiento de los sistemas dinámicos (De Bot, Chan, Lowie, Plat, & Verspoor, 2012).

Los sistemas dinámicos manifiestan características y comportamientos radicalmente distintos a los de los sistemas simples. A diferencia de los sistemas simples, los elementos de los sistemas dinámicos tienen relaciones de interdependencia y codeterminación (García, 2006). En el caso de una casa (ejemplo de sistema simple), realizar cambios a los grifos, tuberías o cables eléctricos no generará cambios en cosas como el número de tabiques en el muro o el color del suelo. Por otro lado, un cambio de magnitud del viento en un determinado clima puede repercutir en humedad y presión ambiental, lo cual afectará la temperatura y con ello a la flora de la región. Este ejemplo del clima muestra cómo las relaciones en los sistemas complejos dificultan, cuando menos, la visión tradicional de elementos de un sistema como variables independientes. Al hablar de sistemas complejos esta constante codeterminación puede hacer que sea imposible establecer relaciones causales o predicciones (Li & Yorke, 1975).

Durante los procesos de aprendizaje de la lengua, o incluso en el uso constante de ésta, la interacción de elementos en la lengua puede generar un soporte entre diversas estructuras durante dicha interacción o una competencia por recursos (Larsen-Freeman & Ellis, 2009). Ellis (2008) ilustra así esta afirmación: cuando una palabra es usada frecuentemente, es imposible escapar de la erosión semántica y morfo-fonológica. Una palabra erosionada en el nivel semántico y morfo-fonológico es menos prominente y, por consiguiente, su retención demanda más esfuerzo, pues este tipo de palabra presenta una contingencia baja en la asociación de forma y función.

Adicionalmente a esta interconexión y codeterminación en el sistema lingüístico, también se observan en él comportamientos típicos de un sistema complejo. En un estudio efectuado por Paul van Geert y Marijn van Dijk (2002) los investigadores tomaron cincuenta y cinco muestras del repertorio léxico de una niña, para analizar su adquisición de preposiciones; estas pruebas se le practicaron de los 18 a los 33 meses de edad. Durante el tiempo que duró el estudio se observó una tendencia general hacia el incremento en el uso de preposiciones, pero en el día a día se re-

portaban fluctuaciones no lineales. Para ilustrar esta afirmación, la investigación mostró hacia el final que la participante utilizó la siguiente cantidad de preposiciones espaciales, en el siguiente orden: 25, 19, 18, 12, 25, 11, 3 y 17. Estas fluctuaciones en el uso de preposiciones no corresponden a un sistema simple en el que los cambios son lineales.

Tanto las relaciones entre los elementos de la lengua como los cambios que se dan en ésta apuntan hacia un sistema complejo. Debido a que los sistemas complejos por definición nunca se encuentran estáticos (aunque en ciertos momentos pueden hallarse en periodos de alta estabilidad), cualquier análisis debe incorporar el tiempo como un factor esencial y necesario (Geert & Dijk, 2002). Cualquier estudio que no muestre cambios o diferencias a lo largo del tiempo no podría tener como objeto los sistemas dinámicos.

Es importante notar que aunque aquí nos pronunciamos a favor de describir el lenguaje como un sistema dinámico, eso no niega que ciertos niveles del lenguaje se puedan describir como sistemas simples. Describir la lengua, o algunos de sus ámbitos como un sistema simple no invalida la interpretación de ésta como un sistema complejo y viceversa. Son dos niveles de estudio que poseen diferentes características.

Para ejemplificar los dos niveles de estudio representados por uno y otro tipo de sistema, podemos hablar en términos de morfología sobre las flexiones verbales. En un sistema simple es posible describir los elementos morfológicos independientes de un verbo (relaciones estables y elementos más o menos autónomos). Incluso, a modo de sistema simple, constatamos cómo estas flexiones pueden ser aplicadas a no-palabras: yo grajupeo, tú grajupeas, nosotros grajupeamos [...]. En contraste, un estudio con la misma temática pero basado en un sistema dinámico, como el de Geert & Dijk (2002), mostraría variabilidad no incremental ni lineal en el uso, acertado o no, de los morfemas verbales y los errores. Un estudio no invalida al otro, ya que por su naturaleza, los sistemas dinámicos tales

como la lengua se encuentran estratificados. En un estrato del sistema, ciertas interacciones tendrán una importancia alta, mientras que en otro, no serán tan importantes (García, 2006). Esta estratificación permite que ciertas relaciones dentro de un nivel puedan ser descritas como simples, mientras que el análisis de uno o varios estratos completos den como resultado relaciones complejas.

Este artículo y su propuesta describen la lengua como un sistema dinámico. Siguiendo a García (2006), se hace una estratificación en una parte de este sistema. En lo que aquí corresponde, esta estratificación del sistema lingüístico es en el nivel léxico. No se intenta sugerir que la lengua sea únicamente un agregado de palabras. Simplemente se trata de un nivel de estudio, como lo han hecho previamente Caspi (2010) o Spoelman & Verspoor (2010). Tal como proceden estos autores, se presentan propuestas metodológicas y prácticas alineadas con conceptos de la TSD: variabilidad, progresión no lineal, estados atractores y especial atención a un estado inicial. De este modo, la propuesta de análisis léxico muestra comportamientos típicos de un sistema dinámico.

Es importante hacer esta aclaración y poner énfasis en los tipos de sistemas debido a que durante casi 400 años se ha mantenido de forma hegemónica una visión de cómo estudiar la lengua (Beniger & Robyn, 1978; García, 2006). Con ello, se advierte un único tipo de construcción mental para describir sus resultados. Es esperable que, con el dominio del método científico y de las representaciones que de ella derivan, exista cierta coerción para mantener el paradigma actual (Kuhn, 1970). En este caso, el paradigma vigente describe la lengua como un sistema simple y no presenta explicaciones satisfactorias con respecto a los cambios no lineales durante la adquisición y el uso de la lengua, adscribiéndola como "ruido" (errores de medición o información irrelevante). Incorporar estudios de sistemas dinámicos (la TSD, por ejemplo) posibilita dar respuestas más adecuadas con respecto a este fenómeno

y otros que competen a este tipo de sistemas. Así pues, para reflexionar sobre patrones comunes en investigaciones léxicas nos planteamos las siguientes preguntas:

- ¿Qué información y datos se pierden al seguir metodologías que intentan reducir el *ruido* y la *variabilidad*?
- ¿Cuáles son las limitantes al representar cambios con valores numéricos a lo largo del tiempo y cómo suplir estos cambios?

En suma, es necesario replantearnos la forma en la que pensamos y representamos la variabilidad y la realidad. No hay una valoración real de qué desconocemos al desechar la variabilidad como *ruido*, ese zumbido molesto que no dice nada mientras mezclamos, promediamos y estandarizamos datos en una representación lineal. Volveremos a este tema en la sección 4.

3. Representaciones visuales: de lo simplificado a lo complejo

3.1. Primer nivel de representación: el punto

Posiblemente la manera más sencilla de evaluar el estado del sistema léxico de un individuo sea generar una batería de ítems —cuya selección, nivel de contextualización y mecánica de respuesta dependerá de los intereses del investigador—, efectuar un conteo de aciertos (o realizar un promedio, si la prueba presenta escalas) y reportar resultados por medio de un puntaje final, descrito por un único valor numérico. Sin importar el método utilizado para evaluar, en última instancia habrá invariablemente un conteo de aciertos o puntajes. Al representar este valor, por lo general se usa un punto a una altura determinada. Este primer nivel de representación es el punto en el plano lineal (Figura 1).



Figura 1. Representación puntual de conteo de aciertos.

La información en este tipo de gráfico es simple y sigue dos nociones metafóricas propuestas por Lakoff & Johnsen (2003): un punto localizado en la sección superior de un plano está asociado con un estado positivo y una mayor cantidad (de aciertos), mientras que un punto localizado en la zona inferior se relaciona con características opuestas. Esto facilita comparaciones. Si el punto del participante "A" se encuentra más elevado que el punto del participante "B", entonces el participante "A" tiene un mejor dominio o conocimiento léxico.

Ya que este tipo de gráfico describe y representa únicamente un puntaje total al comparar participantes, existen elementos no considerados: ¿cuáles y cuántos aciertos o errores comparten los participantes? o ¿es significante la relación puntajes semejantes-ítems compartidos? Es imposible dar respuesta a estas preguntas con estas metodologías y representaciones que predominan en la esfera científica. Estas interrogantes que influyen sobre resultados en el nivel de los ítems aumentan conforme más individuos participan en la generación de un promedio grupal. Una desviación estándar insignificante o casi nula supone resultados altamente homogéneos o consistentes. Como veremos más adelante, esto no es totalmente cierto cuando se toman en cuenta diferencias en el nivel de los ítems.

Para las representaciones puntuales como ésta, el tiempo es un factor no considerado. Sin decirlo explícitamente, estas representaciones sugieren que una única toma de datos léxicos es lo suficientemente representativa como para ilustrar el nivel de adquisición de un elemento alrededor de la toma de la muestra. Estudios longitudinales individuales, en especial aquellos apoyados en la TSD muestran que sobre todo durante el aprendizaje, se suscitan cambios caóticos y muchas veces impredecibles (Caspi, 2010; De Bot, Chan, & Lowie, 2012; De Bot, Lowie & Verspoor, 2007; Rodríguez, 2012). En este tipo de estudios, con facilidad se podría dar una imagen imprecisa del sistema si sólo se tomara una muestra. Esto es particularmente cierto durante los procesos de aprendizaje. El factor tiempo es necesario y, por tanto, constituye un segundo nivel de representación.

3.2. Segundo nivel: el punto en movimiento

El segundo nivel de representación retoma elementos de la representación puntual de la sección anterior (donde el eje vertical representa aciertos), con una adición: el tránsito de izquierda a derecha es el avance en el tiempo. Si un punto representa un momento en el sistema lingüístico, el cambio del sistema lingüístico es únicamente un punto moviéndose, tal como puede verse en la Figura 2.

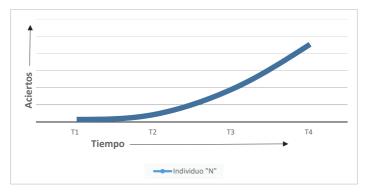


Figura 2. Diseño prototípico de un estudio léxico longitudinal.

La Figura 2 muestra un aprendizaje estereotípico constante y "fluido", donde el punto avanza en el eje del tiempo y sube acorde a su puntaje siguiendo una curva matemáticamente predecible. Esta es una imagen distorsionada de la realidad y de los datos. Los estudios longitudinales que abrazan la TSD subrayan que estas curvas provienen de datos previamente "suavizados" por promedios, proyecciones lineales, interpolación de datos y diferentes tipos de medias (Caspi, 2010). Este suavizado de los datos se hace desechando las irregularidades y la variabilidad para mostrar patrones predecibles. El suavizado de datos ha sido particularmente aprovechado en el ámbito de la lingüística computacional (Jurafsky & Martin, 2008) y la minería de textos (Manning, Raghavan & Schütze, 2009). Estas formas de simplificación permiten mecanizar más eficientemente los datos (y con ellos la realidad), pero no son la realidad. Son hechos interpretados convertidos en constructos teóricos (Feyerabend, 1986). Las mediciones longitudinales que representan de forma menos distorsionada los procesos de aprendizaje léxico muestran avances, retrocesos, y cambios no lineales (Larsen-Freeman, 1997; Larsen-Freeman & Cameron, 2012), como es visible en los datos hipotéticos de puntajes de la Figura 3.

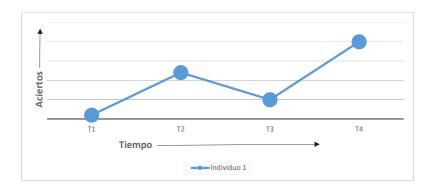


Figura 3. Gráfico prototípicamente esperado para un estudio longitudinal en la TSD.

Aunque la Figura 3 es una representación más realista de los cambios no lineales/complejos durante el aprendizaje del léxico (Beckner, Blythe, Ellis, & Bybee, 2009; Spoelman & Verspoor, 2009), la representación tiene limitantes que se originan desde el conteo de aciertos. Estas limitantes concitan a repensar cómo entendemos el lenguaje. Baste decir al momento que, a diferencia de contar dinero, actividad en la que las monedas son intercambiables si el valor final es el mismo, al medir el crecimiento léxico no deberíamos considerar como intercambiables las palabras aprendidas.

3.3. La recta: el horizonte impenetrable

Supóngase que a un individuo se le aplica tres veces una misma prueba de inventario léxico consistente en 20 ítems, calificándola de forma binaria, con aciertos y errores. En esta prueba, el participante hipotético obtiene en cada una de las tres pruebas seis aciertos de 20 posibles (Figura 4).



Figura 4. Caso hipotético de aparente estabilidad en el aprendizaje léxico.

Al observar la representación y los datos, uno supondría que se ha llegado a un periodo de estatismo en su crecimiento léxico, como lo atestiguan estos totales. Sin embargo, esta aparente estabilidad puede esconder un alto proceso de cambio. Es posible que exista una estabilidad en el nivel superficial (totales), mientras que en el nivel profundo (ítem) haya diferencias significativas, como se propone en la Tabla 1.

Tabla 1 Resultados hipotéticos de una tabla binaria

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11	Ítem 12	Ítem 13	Ítem 14	Ítem 15	Ítem 16	Ítem 17	Ítem 18	Ítem 19	Ítem 20
T1		Х	Χ				Χ					Χ				Χ				Х
T2				Χ			Χ		Χ						Х				Χ	Х
Т3		Χ							Χ				Χ	Х		Χ				Х

Respetando el ejemplo propuesto de una misma prueba de 20 ítems aplicada tres veces, con seis aciertos en cada aplicación, la aparente estabilidad queda en entredicho. Es posible encontrar un fuerte estado de cambio, donde ítems aparentemente conocidos pueden no reaparecer como conocidos o que reaparezcan como conocidos tras un periodo de desconocimiento. La absoluta estabilidad de la línea recta descrita en el plano cartesiano queda comprometida al mirar a detalle la evolución por ítem. Este cuestionamiento de la aparente estabilidad crece considerablemente si en vez de resultados binarios se reportan como resultados escalas numéricas para cada ítem.

3.4. Un nuevo horizonte

Supóngase que a un participante se le aplica tres veces una prueba de reconocimiento léxico de 20 ítems. Cada reactivo se

califica del 0 al 4, siendo 0 el puntaje más bajo y 4 el más alto. En esta prueba se obtienen los promedios T1=1.95, T2=2.1 y T3=2.25 (Figura 5).

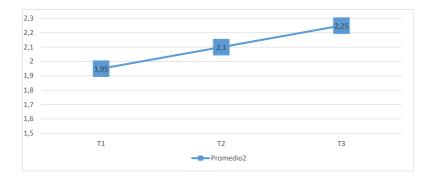


Figura 5. Promedios hipotéticos en la aplicación de una prueba en tres tiempos.

Aunque los promedios muestran una tendencia unidireccional casi lineal hacia la mejoría, en el nivel de los ítems la realidad se muestra menos simple (Tabla 2).

Tabla 2 Ejemplo hipotético de variabilidad compleja

	Ítem 1	Ítem 2	ítem 3	Ítem 4	ítem 5	ítem 6	Ítem 7	ítem 8	ítem 9	ítem 10	Ítem 11	Ítem 12	Ítem 13	Ítem 14	Ítem 15	Ítem 16	Ítem 17	Ítem 18	Ítem 19	Ítem 20	Promedio
T1	0	4	1	0	0	4	4	4	0	0	1	3	4	4	0	2	1	2	1	4	1.95
T2	1	4	1	1	3	0	2	1	0	0	1	3	4	2	2	4	3	3	3	4	2.1
Т3	3	4	1	4	0	0	4	0	0	0	1	3	4	2	2	3	3	4	3	4	2.25

En este ejemplo, los ítems no siguen esta tendencia tan simple. El ítem 14 presenta en la segunda aplicación un resultado "peor" ("2") que en la primera ("4"), mientras que en el ítem 5 se aprecian dos "0" interrumpidos con un "3" en T2, por citar un par de ejemplos. El comportamiento individual de los ítems presenta más patrones que en el promedio: no únicamente se observan tendencias hacia un mayor puntaje (ítem 18); también vemos estabilidades (ítem 3), mejorías seguidas de retrocesos (ítem 5) y retrocesos (ítem 8); en suma, complejidad. Al intentar describir los resultados en el nivel de los ítems mediante la representación de tipo variabilidad versus tiempo, es esperable que los resultados sean de difícil legibilidad (Figura 6).

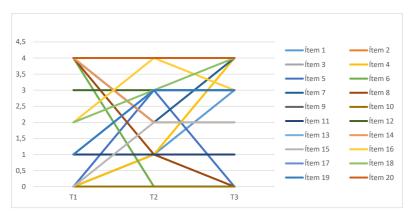


Figura 6. Resultados de una prueba de reconocimiento léxico en tres tiempos por ítem.

En esta prueba hipotética de tan solo 20 ítems, la Fig. 6 entorpece la lectura de los datos en el nivel de los ítems: no perfila tendencias, agrupaciones principales y/o secundarias, relaciones entre ítems y diferencias o semejanzas. Esto es comprensible porque los gráficos en los que el eje "X" representa el tiempo y el eje "Y" la variabilidad, datan de 1600 (Beniger & Robyn, 1978)

y fueron diseñados considerando primordialmente relaciones causales propias de los sistemas simples. Explica Hirschheim (1992) que la ciencia moderna —entendida como el método científico— y por ende esta representación de la realidad, describe qué sucede en el mundo social al buscar regularidades y relaciones causales entre sus elementos constituyentes. Este proceso se basa necesariamente en el reduccionismo, para el cual todo es segmentado de manera progresiva en constituyentes y relaciones más elementales. Resulta entonces un verdadero contrasentido intentar plasmar comportamientos complejos con herramientas que están diseñadas en específico para simplificar y reducir las interacciones de un sistema. Ilustrar sistemas complejos con herramientas creadas dos siglos antes (Holmes, 2008) muestra los peligros latentes cuando los investigadores se asumen como usuarios de herramientas, más que como creadores de nuevas herramientas, pues "cuando tu única herramienta es un martillo en mano, el mundo parece un clavo" (Hirschheim, 1992). Este comentario es, para nosotros, ejemplo del compromiso necesario para generar herramientas y representaciones que nazcan con la complejidad en su esencia, no simplemente forzar otras sólo porque son parte de una cultura dominante.

3.5. Trazar mapas de la complejidad

En la sección anterior se presentó el ejemplo hipotético de una prueba de conocimiento léxico aplicada tres veces, en la cual se hacía uso de una escala del 0 al 4. Tradicionalmente se graficaría en un plano de variabilidad *versus* tiempo, pero es posible crear un híbrido en el que variabilidad y tiempo se fusionen en ambos ejes del plano cartesiano (Ibarra, 2017).

En el eje Y, la puntuación de T3 se anota de forma ascendente (0, 1, 2, 3, 4). El eje X contiene las puntuaciones de T1 y T2 concatenadas, es decir, una adyacente a la otra para generar un código numérico de dos posiciones que inicia con la puntuación

de T1 y termina con la de T2. Por ejemplo, un resultado T1=2 y T2=1 produciría el código 21. Por tanto, el eje X parte desde 00 y sigue con 01, 02, 03, 04, 10, 11, 12, 13, 14, hasta llegar a 44 (no se muestra en la Fig. 8). De esta manera, el eje X cubre todas las posibles combinaciones de resultados en T1 y T2. Al cruzar el eje X (que contiene T1 y T2) con el eje Y (que contiene T3), podemos graficar en una localización estos puntajes. En esta propuesta, cada posible evolución de un ítem es una posición en plano cartesiano. Graficar es, entonces, crear un agregado de puntos en el plano.

Si, por ejemplo, un ítem presenta los resultados: T1=0, T2=1 & T3=4; entonces en el plano cartesiano el eje X=T1 T2=01; mientras que el eje Y=T3=4 (7).

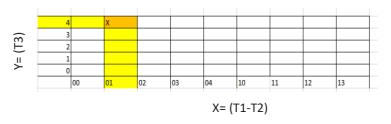


Figura 7. Ejemplo de gráfico de un punto.

El mapa resultante contiene todas las posibles combinaciones de puntuaciones a lo largo de las tres pruebas. Su organización es tal que puntos contiguos presentan el menor cambio posible (únicamente de un punto en T2 o T3). Esto implica que aquellos ítems que se hallan agrupados como racimos presentan puntajes muy parecidos, mientras que dos distanciados tienen comportamientos diferentes. Ya que este mapa está ordenado con valores numéricos por eje, ciertas zonas muestran mejoría en puntuaciones, mientras que otras, disminución en los puntajes (Figura 8).

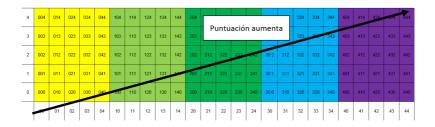


Figura 8. Mapa dinámico y sus constituyentes.

En la Figura 8 se aprecian diferentes bandas de colores. Esto es porque aunque las casillas contiguas presentan un cambio de un punto en T2 y T3; T1 funge como denominador común en forma de bandas. En buena medida, esto es porque bajo los lineamientos de la TSD, una característica esencial es que los sistemas son extremadamente sensibles a condiciones iniciales (Oestreicher, 2007; De Bot, Lowie, & Verspoor, 2007).

En pruebas de múltiples reactivos, es esperable que ciertos ítems tengan puntuaciones semejantes. En estos casos se coloca en la casilla cuántos ítems presentan un mismo comportamiento. Cuando un mapa tenga ítems disgregados, esto será indicio de un sistema en reordenamiento (De Bot, Lowie, Thorne, & Verspoor, 2013; García, 2006), incluso cuando el promedio pudiera sugerir total estabilidad. Un mapa con ítems concentrados en uno o varios puntos es sintomático de estabilidad o de cercanía a ésta. Cuando los ítems orbitan alrededor de coordenadas estables (000, 111, 222, 333, 444) esto sugiere que el sistema lingüístico se está asentando en estados atractores (Plaza-Pust, 2008), ya sea hacia el aprendizaje o hacia la fosilización.

Al graficar el ejemplo de tres tiempos comentado en secciones anteriores (Tabla 2), las respuestas aparecen dispersas a lo largo del mapa, sugiriendo un sistema caótico. Aunque hay algunos ítems que tienen comportamientos idénticos (sugerido por casillas con puntuación superior a 1), las respuestas no

parecen concentrarse en una zona particular del mapa, lo que sugiere un comportamiento que aún se encuentra en cambio (Figura 9).

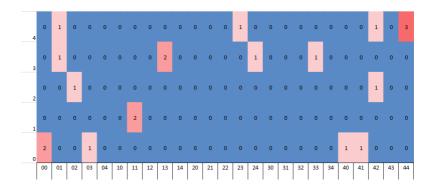


Figura 9. Ejemplo hipotético de mapa dinámico.

Como ejercicio mental —y para mostrar la importancia de los cambios en el nivel de los ítems—, se reordenarán únicamente los resultados por ítem en cada una de las pruebas, yendo de menor a mayor. Esto genera una mayor estabilidad en el nivel de los ítems pero con promedios idénticos (Tabla 3).

Tabla 3
Ejemplo hipotético ordenado

		Ítem 1	Ítem 2	ítem 3	Ítem 4	ítem 5	ítem 6	Ítem 7	ftem 8	ftem 9	ítem 10	Ítem 11	Ítem 12	Ítem 13	Ítem 14	Ítem 15	Ítem 16	Ítem 17	Ítem 18	ftem 19	ítem 20	Promedio
	T1	0	0	0	0	0	0	1	1	1	1	2	2	3	4	4	4	4	4	4	4	1.95
ſ	T2	0	0	0	1	1	1	1	1	2	2	2	3	3	3	3	3	4	4	4	4	2.1
ſ	Т3	0	0	0	0	0	1	1	2	2	3	3	3	3	3	4	4	4	4	4	4	2.25

En este caso, el mapa (o gráfica) muestra puntuaciones agrupadas alrededor de los estados atractores o sitios con puntuaciones idénticas (Figura 10).

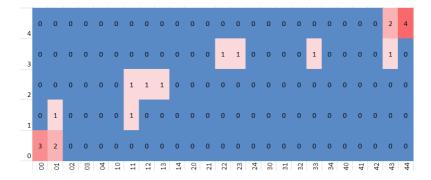


Figura 10. Mapa dinámico con datos hipotéticos organizados.

Las diferencias entre los mapas dinámicos aquí presentados (Figuras 9 y 10) muestran algunas limitantes de formas más tradicionales de ilustrar e investigar la lengua. Ciertamente es un punto de vista diferente, pero ¿cuáles son sus implicaciones?

4. Retomando la complejidad desde el horizonte

Hasta aquí hemos explicado y ejemplificado cómo, a menudo, bajo el método científico, se privilegian las simplificaciones en aras de mecanizar relaciones. La mecanización permite generar más y mejores predicciones. La utilidad de predecir no puede ser minimizada. Como mencionamos anteriormente, algunas áreas de los estudios del lenguaje, como la lingüística computacional, han explotado en particular la predicción de datos lingüísticos. Las representaciones gráficas que son consistentes con la intención de utilizar predicciones, reducen los datos y sus relaciones a las más importantes. Es innegable la utilidad de este modo de pensar y actuar, pero consideramos indispensable señalar que hacerlo irreflexivamente nos limita.

La forma en la cual se recogen, procesan e interpretan los datos, señalan cuestionamientos y problemáticas interesantes,

que abarcan lo estadístico, lo epistemológico y lo filosófico. A continuación presentamos algunas reflexiones acerca de estos problemas.

En cuanto a la medición estadística, pocos discrepan con la visión de que las medidas de tendencia central por lo común usadas al describir el crecimiento léxico son superficiales, a falta de un mejor adjetivo para describirlas. Lo son de la misma forma que la superficie del agua no informa de las corrientes del mar: se establece un borde limítrofe conveniente que, sin embargo, oscurece procesos tales como cambios y movimientos internos. Las medidas de tendencia central, conteos totales y promedios ignoran todo lo que no sea descriptible en un único valor numérico por observación. Aunque las medidas de tendencia son complementadas con medidas de dispersión, estas últimas usualmente sirven a lo mucho como un agregado que no aporta información relevante.

En el caso del aprendizaje de lenguas, podemos imaginar algún contexto en el cual podría ser útil y práctico reducir el conocimiento del vocabulario de una lengua a un valor numérico (como en la Tabla 2); por ejemplo, al sugerir que alguien tiene el dominio del 70% de ésta. De hecho, este tipo de afirmaciones resultan bastante cotidianas. En realidad, ese es un dato de escaso significado, pues es muy difícil establecer cuál es el tamaño del vocabulario de una lengua, cuál es el tamaño del vocabulario "normal" de un hablante nativo o qué tan bien se conoce el vocabulario. Hablar de conocer un 70% sólo tiene sentido si se compara con un 50% o un 75%, pero no siempre esas medidas son realmente comparables, pues no son realizadas con los mismos instrumentos. El tamaño del vocabulario es sólo una de las dimensiones del conocimiento de éste; otras pueden ser la profundidad, la organización y el acceso a aquél (Meara, 2005). Como plantea Serres (1991: 21), "la simplificación y las elecciones forzadas oscurecen la compleja realidad; esta forma de maniqueísmo no permite un pensamiento nuevo".

Otra de las consecuencias del uso de datos complejos como los que hemos presentado aquí es que se produce el cuestionamiento de un fundamento básico de la medición de lenguas (y de la psicometría en general): la confiabilidad (el otro fundamento es la validez). La confiabilidad es la variabilidad que deben presentar las pruebas. Este principio postula que los resultados de una misma prueba realizada por un mismo sujeto varias veces de manera consecutiva deberían ser prácticamente los mismos. Si los resultados de aplicar varias veces una misma prueba a un mismo participante de manera consecutiva son exactamente iguales, entonces la confiabilidad de la prueba es absoluta. De lo contrario, hay un error de medición, y la prueba no es totalmente confiable.

En los casos que hemos presentado en la sección 3.3. nos centramos no sólo en el comportamiento de los participantes sino en el comportamiento de cada palabra para cada participante. Los resultados obtenidos en investigaciones relacionadas con sistemas dinámicos contradicen siempre, en distinto grado, el principio de confiabilidad test-retest (Ibarra, 2017; Spoelman & Verspoor, 2010; Geert & Dijk, 2002). Es decir, que una misma prueba presentada a los mismos participantes bajo distintos formatos con lapsos de separación de pocas semanas arroja resultados muy diferentes. Rodríguez (2012) reporta casos en los cuales, con lapsos de sólo tres días, se muestra variabilidad en al menos un tercio de los datos obtenidos. Este tipo de hallazgos entra en conflicto directo con una de las piedras angulares de la medición, que, como mencionábamos, rechazaría estos datos como errores de medición de un instrumento poco confiable o "suavizaría" la variabilidad o el ruido promediando o interpolando datos. Esta problematización del concepto de confiabilidad es un tema de intenso y amplio debate en la comunidad académica. Algunos investigadores que promueven el uso de una evaluación dinámica (más cualitativa, distinta de la que hemos aquí expuesto) han llegado a conclusiones muy parecidas a las que estamos exponiendo (Poehner, 2008).

La visualización de datos es una manera de narrar ciertos eventos y su importancia es capital. No son una simple decoración para los datos. Los gráficos poseen, de ser utilizados correctamente, un valor inherentemente explicativo. Menciona Tufte (2001) ejemplos en los cuales diversas catástrofes pudieron haber sido advertidas de haber contado con gráficas que perfilaran la información adecuada de los datos recabados. Un ejemplo es la relación entre la ubicación de la bomba de agua en Londres respecto de los fallecidos durante la epidemia de cólera en 1854. Asimismo, la explosión del Challenger pudo haber sido evitada de haber visto un conjunto de gráficos que mostraran los efectos de las bajas temperaturas ambientales en las turbinas del cohete. Los gráficos, como interfaz narrativa que media entre los datos puros y el lector, deben ser tratados como algo más allá que un añadido a los datos.

Al analizar las gráficas de tipo variabilidad *versus* tiempo, se mostró que es posible que dos estudios de evolución léxica con un mismo puntaje y desviación estándar presentaran niveles distintos de variabilidad en el nivel de los ítems (Tabla 2 y Tabla 3). Bajo la misma línea argumentativa, también se explicó que aunque se mostrara una absoluta estabilidad en puntajes totales, eso no significaría que sus respuestas fuesen totalmente estables. Aún dentro de una total estabilidad, hay espacio para una variabilidad imperceptible al simple conteo de aciertos. Creemos haber mostrado e ilustrado que el aprendizaje no es necesariamente un proceso lineal, constante, progresivo e incremental sino que el aprendizaje es un proceso complejo.

Estamos conscientes de que una medición compleja de los fenómenos como la que estamos proponiendo es sólo factible cuando se utilizan herramientas sofisticadas en la recolección y el procesamiento de datos. Sin embargo, no consideramos que el uso masivo de datos vaya a facilitar un conocimiento absoluto. Al igual que muchos pensadores contemporáneos (Sadin, 2015), no creemos que todo sea mensurable y cuantificable. No consideramos tampoco que el manejo masivo de datos tratados

por un algoritmo consiga mostrar correlaciones secretas que hasta ahora habían permanecido ocultas. No deja de resultar curioso que tanto los planteamientos más tecnófilos, como los planteamientos más *anticientifistas* (Angell & Demetis, 2012) coincidan en que la causalidad es un obstáculo en el desarrollo de un conocimiento de la realidad.

Así, ante la sugerencia de que las correlaciones vayan a sustituir a las causalidades (que el "así son las cosas" sustituya al por qué), coincidimos en la crítica que hace Han (2014) cuando advierte que "la cuantificación de lo real en búsqueda de datos expulsa al espíritu del conocimiento". A pesar de ello, nuestra visión no es anticientifista y sí compartimos la idea de que "frente al orden clásico llegado hasta nosotros y llevado a una potencia jamás conocida en el campo de los métodos estamos obligados a reconocer la realidad del desorden" (Serres, 1991: 53) y que ese desorden puede ser descrito por formas innovadoras y complejas de representación. Dicho de otro modo, consideramos que entre la imposibilidad cartográfica del emperador chino de Borges y la simplificación extrema de una etiqueta numérica, hay propuestas en las que es posible combinar discursos cientifistas con discursos humanistas. Lo que ocurre es que las propuestas para trazar estos espacios complejos (y complicados) a menudo no son bien recibidos en el ambiente de hiperespecialización que impera en la universidad. La TSD y su hincapié en las redes de relaciones complejas en sistemas abiertos es una apuesta atractiva para ese va no tan nuevo derrotero de ciencia.

Finalmente, un cambio de mirada en la medición de datos conllevaría muchas implicaciones. Algunos especialistas en medición del conocimiento de segundas lenguas, como Mislevy, Steinberg & Almond (2002), creen que las últimas investigaciones invitan a la implementación de nuevos constructos para medir el conocimiento de lenguas. Los especialistas tendrán que explicar cómo esos nuevos constructos describen el comportamiento de personas concretas en situaciones específicas. Y lo tendrán que hacer de una manera comprensible para aprendices,

profesores, padres, empleadores y administradores. Contrariamente a lo que sucede con el modo tradicional de evaluar, habría que realizar evaluaciones complejas y multidimensionales, y el éxito que puedan tener estas nuevas propuestas dependerá de la cooperación y el diálogo que haya entre todas las partes involucradas en el aprendizaje.

Conclusión

En este artículo hemos propuesto que, dependiendo de cómo se recojan y procesen los datos, se puede extraer información útil, capaz de impulsar el debate sobre el significado de aquéllos y el cuestionamiento de conocimientos previos que podrían considerarse simplistas. No obstante, queremos insistir en que el uso masivo de datos no puede suplir ni eliminar las explicaciones teóricas. Eso significaría el final de la verdad, el final de la narración. Se requiere que los científicos tomen las riendas de los datos, no que los datos constriñan el actuar de los científicos. Como dice Han (2015), "los datos son meramente aditivos. La adición se opone a la narración. A la verdad le es inherente una verticalidad. Los datos y las informaciones, por el contrario, habitan lo horizontal".

Aunque los gráficos no presentan información propiamente nueva, pues son solamente representaciones visuales derivadas de datos numéricos, su fortaleza radica en dar forma y figura a números en abstracto, generando un concentrado de datos fácilmente comprensible. Sirven, en cierta manera, de filtro para comprender una ventana de la realidad: cierta información se perfila y se lleva al frente, como eje narrativo. Anteriormente se ha abogado en que el filtro está matizado hacia las representaciones lineales y relaciones causales, lo cual puede ser una limitación. Es importante estar consciente de cómo es que estos vínculos entre el intérprete y los datos pueden limitar el pensamiento del primero, para así suponer si existen alternativas capaces de suplir

dichas carencias. En este sentido, creemos en lo adecuado del enfoque de Serres (1991: 63), quien afirma lo siguiente:

Hoy debemos proponer un modelo nuevo para nuestros nuevos problemas. Hay orden en el desorden, hay desorden en el orden. Nuestras redes están inmersas localmente en las nubes, nuestras estructuras en las distribuciones, como archipiélagos en el mar. Pero también hay nubes en las redes, y mar entre las islas. Este modelo es sin embargo demasiado escenográfico, parece aún inmerso en un espacio global del que nada sabemos, es también casi estático. Una vez más debemos meditar sobre el tiempo.

Tenemos 400 años estudiando con base en el punto y el punto en movimiento. Merecemos conocer más, a mayor profundidad y en direcciones que aún no hemos considerado. Intentar agotar temas estudiados no es suficiente. También es necesario reconsiderar los filtros que hacen factible observar la realidad y las representaciones que nos posibilitan entender estos filtros. Tal vez la realidad sea más compleja que lo que ahora entendemos. No podremos saberlo hasta que observemos bajo una nueva lupa y nos adentremos en terrenos inexplorados.

Referencias

- ANGELL, Ian O. & DEMETIS, Dyonisios (2012). *Science's First Mistake: Delusions in Pursuit of Theory*. Nueva York: Bloomsbury Academic. Disponible en: https://sciencesfirstmistake.files.wordpress.com/2010/07/sciences first mistake angelldemetis.pdf
- BECKNER, Clay; BLYTHE, Richard; ELLIS, Nick C. & BYBEE, Joan (2009). "Language is a Complex Adaptive System: Position Paper". *Language Learning*, núm. 59. Supl. 1. Michigan: Universidad de Michigan. Disponible en: http://cnl.psych.cornell.edu/pubs/2009-LACAS-pos-LL.pdf
- Beniger, James Ralph & Robyn, Dorothy (1978). "Quantitative Graphics in Statistics: A Brief History". The American Statisti-

- *cian*, vol. 32, núm. 1, pp. 1–11. Disponible en: https://doi.org/10.1 080/00031305.1978.10479235
- CASPI, Tal (2010). A dynamic perspective on second language development [Tesis de Doctorado]. Groningen: University of Groningen.
- De Bot, Kees; Chan, Huiping; Lowie, Wander; Plat, Rika & Verspoor, Marjolijn (2012). "A Dynamic Perspective on Language Processing and Development". *Dutch Journal of Applied Linguistics*, vol. 1, núm. 2, pp. 188–218. Disponible en: https://doi.org/10.1075/dujal.1.2.03deb
 - ; Lowie, Wander; Thorne, Steven & Verspoor, Marjolijn (2013). Dynamic Systems Theory as a Comprehensive Theory of Second Language Development. En P. García Mayo, J. Gutiérrez Mandago & M. Martínez Adrián (eds.). *Contemporary Perspectives on Second Language Acquisition*, pp. 199–220. Ámsterdam: John Benjamins Publishers.
 - ______; Lowie, Wander & Verspoor, Marjolijn (2007). "A Dynamic Systems Theory Approach to Second Language Acquisition". *Bilinguism: Language and Cognition*, vol. 10, núm. 1, pp. 7–21. Cambridge: Cambridge University Press. Disponible en: https://doi.org/S1366728906002732
- Ellis, Nick C. (2008). "The Dynamics of Second Language Emergence: Cycles of Language Use, Language Change, and Languaje Acquisition". *The Modern Language Journal*, vol. 92, núm. 1, pp. 232-249.
- FEYERABEND, Paul (1986). Tratado contra el método. Esquema de una teoría anarquista del conocimiento. Trad. de Diego Ribes. Tecnos: Madrid.
- Freire, Paulo (1970). "The Adult Literacy Process as Cultural Action for Freedom". *Harvard Educational Review*, vol. 40, núm. 2, pp. 205-225.
- GARCÍA, Rolando (2006). Sistemas complejos. Conceptos, método y fundamentación epistemológica de la investigación interdisciplinaria. Barcelona: Gedisa.

- GEERT, Paul van & DIJK, Marijn van (2002). "Focus on Variability, New Tools to Study Intra-individual Variability in Developmental Data". *Infant Behavior and Developments*, vol. 25, núm. 4, pp. 340–374.
- Habermas, Jürgen (2002). Ciencia y técnica como "ideología". Trad. de Manuel Jiménez Redondo. Madrid: Tecnos.
- HAN, Byung-Chul (2014). *Psicopolítica. Neoliberalismo y nuevas técnicas de poder.* Trad. de Alfredo Bergés. Barcelona: Herder Editorial S.L.
- _____ (2015). *La salvación de lo bello*. Barcelona: Herder Editorial S.L.
- HIRSCHHEIM, Rudy A. (1992). Information Systems Epistemology: An Historical Perspective. En R. Galliers (ed.), *Information Systems Research: Issues, Methods and Practical Guidelines*. Londres: Blackweel Scientific Publications.
- Holmes, Philip (2008). *A Short History of Dynamical Systems Theory:* 1885-2007. En Encyclopedia of Life Support Systems: Mathematical Science. Oxford: UNESCO-EOLSS. Disponible en: http://www.eolss.net/sample-chapters/c02/e6-132-21.pdf
- IBARRA GONZÁLEZ, Rodrigo (2017). La teoría de sistemas dinámicos aplicada al desarrollo léxico de cognados y falsos cognados: una propuesta metodológica [Tesis de Maestría]. Universidad Autónoma de Querétaro.
- JURAFSKY, Daniel & MARTIN, James H. (2008). Speech and Language Processing: An Introduction to Language Natural Processing, Computational Linguistics, and Speech Recognition. Upper-Saddle River, Nueva Jersey: Pearson/Prentice Hall.
- Kuhn, Thomas S. (1970). *The Structure of Scientific Revolutions*. Chicago: Universidad de Chicago.
- LAKOFF, G. & Johnsen, M. (2003). *Metaphors we live by*. Chicago: University of Chicago Press.

- Larsen-Freeman, Diane (1997). "Chaos/Complexity Science and Second Language Acquisition". *Applied Linguistics*, vol. 18, núm. 2, pp. 141-165. Disponible en: https://www.uibk.ac.at/anglistik/staff/freeman/coursedocuments/diane chaos paper.pdf
- & Cameron, Lynne (2012). "Complex Systems and Applied Linguistics". *International Journal of Applied Linguistics*, vol. 17, núm. 2, pp. 226-239. Oxford: Oxford University Press.
- & Ellis, Nick C. (2009). Language as a Complex Adaptive System. Michigan: John Wiley & Sons.
- LI, Tien-Yien & YORKE, James A. (December 1975). "Period Three Implies Chaos". *The American Mathematical Monthly*, vol. 82, núm. 10, pp. 985-992. Disponible en: http://www.uvm.edu/~cdanfort/courses/266/period-3.pdf
- Manning, Christopher D., Raghavan, Prabnakar & Schütze, Hinrich (2009). *An Introduction to Information Retrieval*. Cambridge: Cambridge University Press.
- MEARA, Paul (2005). Designing Vocabulary Tests for English, Spanish and Other Languages. En C. Butler, María A. Gómez-González & Susana María Doval Suárez (eds.). *The Dynamics of Language Use. Functional and Contrastive Perspectives*. Ámsterdam: John Benjamins Publishing Company.
- MISLEVY, Robert J.; STEINBERG, Linda S. & ALMOND, Russell G. (2002). "Design and Analysis in Task-Based Language Assessment". *Language Testing*, vol. 19, núm. 4, pp. 477–496. Disponible en: https://doi.org-/10.1191-/0265532202lt2410a
- OESTREICHER, Christian (September 2007). "A History of Chaos Theory". *Dialogues in Clinical Neuroscience*, vol. 9, núm. 3, pp. 279–289.
- PLAZA-Pust, Carolina (Summer 2008). "Dynamic Systems Theory and Universal Grammar: Holding up a Turbulent Mirror to Development in Grammars". *Modern Language Journal*, vol. 92, núm. 2, pp. 250–269.

- POEHNER, Matthew E. (2008). Dynamic Assessment. A Vygotskian Approach to Understanding and Promoting L2 Development. Berlín: Springer Publishing.
- RODRÍGUEZ SÁNCHEZ, Ignacio (2012). "La teoría de los sistemas dinámicos como marco para el estudio de la adquisición de vocabulario de L2. Tres estudios empíricos". En Amalia Hoyos Arvizu y Ana Lidia Munguía (coord.). *Interdisciplinas Lingüísticas y Lingüística Aplicada*, vol. 5, pp. 143–159. Sonora: Universidad de Sonora.
- SADIN, Éric (2015). La Vie Algorithmique: Critique de la Raison Numérique. París: L'Échappée, coll. Pour en finir avec.
- SERRES, Michel (1991). *El Paso del Noroeste: Hermes V.* Barcelona: Debate, col. Pensamientos.
- Spoelman, Marianne & Verspoor, Marjolijn (2010). "Dynamic Patterns in Development of Accuracy and Complexity: A Longitudinal Case Study in the Acquisition of Finnish". *Applied Linguistics*, vol. 31, núm. 4, pp. 532–553. Disponible en: https://doi.org/10.1093/applin/amq001
- Tufte, Edward R. (2001). *The Visual Display of Quantitative Information*. Cheshire, Connecticut: Graphics Press. Disponible en: https://www.colorado.edu-/geography/foote/maps/assign/reading/TufteCoversheet.pdf
- VYGOTSKY, Lev (1978). Interaction Between Learning and Development. En M. Cole, V. John-Steiner, S. Scribner & E. Souberman (eds.). *Mind and Society: The Development of Higher Psychological Processes*. Cambridge, MA.: Harvard University Press. Disponible en: http://www.psy.cmu.edu-/~siegler/vygotsky78.pdf