

Repositorio abierto de locuciones de fórmulas matemáticas

M Antonia Huertas¹, Mireia Pascual³, César Córcoles¹, Laia Llorens² y Roger Griset²

¹Estudios de Informática, Multimedia y Telecomunicación

²Gestión de Contenidos

³Servicios Audiovisuales

Universitat Oberta de Catalunya.

Rambla del Poblenou, 156, 08018 Barcelona.

{ccorcoles, mhuertass, ellorensp, mpascual, rgriset}@uoc.edu

Abstract. El aprendizaje de la lectoescritura de fórmulas y expresiones matemáticas en un entorno de aprendizaje presencial recae en el profesorado que enseña a leerlas de forma correcta. En un entorno virtual de aprendizaje no hay un elemento que sustituya al profesor hablante ausente. Aprovechando las posibilidades tecnológicas para digitalizar la voz humana y posteriormente adjuntar los archivos de sonido a textos web un equipo formado por profesores de matemáticas y técnicos de gestión de materiales web y de sonido hemos diseñado e implementado una prueba piloto consistente en un primer repositorio de locuciones de fórmulas matemáticas utilizadas en el aprendizaje de un módulo didáctico de matemáticas para estudiantes de Ingeniería Informática. En este artículo explicamos y evaluamos la prueba y después presentamos el proyecto que es su continuación: un repositorio abierto básico, multilingüe y amplio de locuciones de fórmulas matemáticas.

Keywords: repositorio, contenido abierto, lectoescritura matemática, MathML, taxonomía, ontología.

1 Introducción

La revolución digital del siglo XX ha alcanzado ya de pleno los entornos tradicionales de enseñanza y aprendizaje. Paralelamente a la aparición de las universidades totalmente en línea, los entornos de aprendizaje virtuales, sus tecnologías y prácticas pedagógicas están modificando todos los tipos de enseñanza y aprendizaje, presencial y no presencial, formal e informal. Aunque la tecnología de la información y la

2 M Antonia Huertas¹, Mireia Pascual³, César Córcoles¹, Laia Llorens^{P2P} y Roger Griset²

comunicación está impregnando rápidamente las instituciones educativas, y nuevos tipos de estudiantes llamados nativos digitales están ya en nuestras aulas, los usos de esas tecnologías siguen limitados por las prácticas educativas y las creencias de profesores formados en la era pre-e-learning.

En el caso de un entorno de aprendizaje no presencial como es una universidad a distancia totalmente en línea el desencuentro entre la tecnología disponible y las necesidades educativas resulta paradójico. Es el caso de la experiencia que presentamos en este artículo.

El contexto del caso que trataremos es la enseñanza y aprendizaje de matemáticas en un entorno virtual de aprendizaje no presencial. En este contexto las habituales dificultades para la enseñanza de las matemáticas se ven aumentadas y aparecen otras nuevas. Uno de estos nuevos problemas es el del aprendizaje de la forma correcta de verbalizar las fórmulas y las expresiones matemáticas. En la enseñanza presencial la expresión escrita y oral de las matemáticas (tanto expresiones típicas como fórmulas) recae sobre el profesor, que es el encargado de la transmisión oral y escrita del lenguaje matemático. Los textos de matemáticas, por otro lado, sólo se ocupan de la introducción a la escritura correcta de las expresiones y fórmulas matemáticas y presuponen que los y las lectoras ya conocen la forma correcta de verbalizarlas. ¿Quién nos enseñó a decir “tres más dos igual a cuatro” ante la fórmula $3+2 = 4$? ¿Cómo hemos aprendido la forma de leer una fórmula como $\cos^2(x) + \sin^2(x) = 1$? ¿Recordamos cuándo aprendimos a escribir con fórmulas que el veinte por ciento de una cantidad equivale a su quinta parte? La enseñanza formal de esta parte de las matemáticas, que llamaremos lectoescritura de fórmulas y expresiones, ha quedado formalmente oculta por la figura siempre presente del profesorado.

En la enseñanza no presencial ese profesor o profesora no transmite oralmente su conocimiento. Y, sin embargo, en los textos y recursos que se utilizan en la enseñanza virtual nada ni nadie se encarga de transmitir esa lectoescritura. En la enseñanza a distancia, incluso en el caso de los cursos en línea que tienen acceso a las tecnologías audiovisuales digitales, la transmisión verbal de conocimiento ha desaparecido, ya que los hipertextos y el *software* matemáticos son normalmente textos escritos. La presuposición implícita en los textos y manuales de matemáticas de que la lectura de las fórmulas las enseña el profesor en la clase se ha mantenido invisible al pasar a la enseñanza-aprendizaje no presencial y a usar hipertextos y textos con contenido matemático sin que ningún elemento nuevo haya sustituido a la figura del profesor hablante. La situación común en un entorno no presencial es, por tanto, que el estudiante ve la fórmula escrita, la puede interpretar pero no sabe como se lee o por lo menos no tiene seguridad en cómo se lee.

No son de extrañar situaciones como la de una estudiante de Álgebra que en la única sesión presencial de nuestra universidad virtual preguntó qué era esa A invertida: era el símbolo matemático para el cuantificador universal “para todo” (\forall), o la manera de referirse al símbolo \int como el “gusano” en lugar de “integral”.

A partir del conocimiento de esa necesidad en nuestros estudiantes, y de las posibilidades tecnológicas para digitalizar la voz humana, que según Mayer [2] es mejor para aprender que una voz mecánica, y posteriormente adjuntar los archivos de sonido a textos web, un equipo formado por profesores de matemáticas y técnicos de gestión de materiales web y de sonido diseñamos la prueba piloto que os presentamos a continuación. Se trata de un primer repositorio de locuciones de fórmulas matemáticas que han sido utilizadas en el aprendizaje de un módulo didáctico de matemáticas para estudiantes de ingeniería informática. Discutiremos la evaluación de la prueba y presentaremos el nuevo diseño del proyecto que es su continuación: un repositorio abierto básico y amplio de locuciones de fórmulas matemáticas.

2 Prueba piloto de locución de fórmulas matemáticas

Mediante el uso de técnicas multimedia y del nuevo lenguaje MathML específico para la edición de fórmulas matemáticas en la web, se ha remodelado una parte de los materiales didácticos de la asignatura de Análisis Matemático de la Ingeniería Técnica Informática. Concretamente se rediseñó e implementó el módulo de Funciones Complejas que es un contenido totalmente nuevo ya que sus conceptos y símbolos matemáticos no se han estudiado previamente ni en secundaria, ni en otras asignaturas de la carrera. Por ello, en esta prueba nos asegurábamos que muchas de las fórmulas que aparecerían serían nuevas para los estudiantes y que seguramente no sabían a priori como se leían.

El formato del material era un archivo XHTML-MathML accesible con cualquier navegador de uso común (utilizando, en algún caso, *plug-ins* de acceso gratuito). Las locuciones se habían realizado previamente en un trabajo conjunto entre una profesora de matemáticas y una técnica de audiovisuales. Las locuciones se registraron en formato mp3 con una resolución adecuada para su distribución a través de la web (lo que impide el uso de grandes resoluciones) pero conservando la claridad necesaria para su fácil comprensión. Para la inclusión en los diferentes archivos XHTML se utilizó una API JavaScript de libre uso y distribución, SoundManager2¹, que hace uso de la tecnología Flash para reproducir bajo demanda los sonidos. En una primera fase, los archivos únicamente se etiquetaron por su orden de aparición en el material y se almacenaron en una carpeta específica del proyecto web.

El objetivo de la prueba era investigar si las locuciones de fórmulas servían efectivamente para que los estudiantes aprendieran a leer esos nuevos símbolos y conocer la satisfacción de los mismos respecto al nuevo material.

El método utilizado para la investigación consta de dos partes:

En la primera parte se hizo un taller de lectura presencial al inicio del semestre, con estudiantes que aún no habían tenido contacto con el material web con locuciones. El objetivo del taller era conocer el nivel de lectura de fórmulas matemáticas de nivel

¹ <http://www.schillmania.com/projects/soundmanager2/>

pre-universitario. Para ello se les pasó un cuestionario (C1) donde se pedía a los estudiantes transcribir tal cual leyeran 15 fórmulas y expresiones matemáticas. Una vez hubieron contestado el cuestionario se realizó el taller, donde una profesora explicó la forma de leer cada una de las fórmulas del cuestionario y presentó una tabla con la transcripción de la lectura de dichas fórmulas. Seguidamente se entregó a los mismos estudiantes un nuevo cuestionario (C2) con 15 fórmulas de tipología y dificultad semejante al primer cuestionario para que transcribieran de nuevo la lectura de las fórmulas, pero ahora podían utilizar la tabla de lectura presentada. Los cuestionarios fueron recogidos y evaluados.

Resumiendo, la experiencia consistió en pasar dos cuestionarios, C1 y C2, a 15 estudiantes que asistieron voluntariamente al taller de lectura.

- El cuestionario C1 es un test de 15 preguntas, el estudiante tiene que leer una fórmula sin ayuda de tablas de lectura
- El C2 es un test de 15 preguntas, el estudiante tiene que leer una fórmula con ayuda de tablas de lectura que han sido presentadas y explicadas por la profesora.

El porcentaje de aciertos de cada uno de los 15 estudiantes en cada uno de los cuestionarios (C1,C2) queda reflejado en la tabla 1, así como el año en el que el estudiante había seguido un curso de matemáticas por última vez. En dicha tabla se han identificado dos grupos, el de los estudiantes que hacía más de dos años que no seguían un curso de matemáticas (en los cuatro casos hacía más de diez años), y el de los que hacía dos años o menos que habían seguido un curso de matemáticas (éstos últimos coinciden con los que el último curso seguido había sido en la UOC y por tanto no presencial). En ambos grupos se ha calculado el promedio de aciertos, antes y después de realizar el taller y usar las tablas.

Table 1. Resultados de los dos cuestionarios del taller de lectoescritura presencial según el último año en que se había seguido un curso de matemáticas.

Último año	C1-% aciertos	C2-% aciertos
1979	67%	80%
1987	47%	47%
1992	20%	13%
1994	53%	100%
Promedio	47%	60%
2005*	93%	100%
2006*	53%	53%
2006*	87%	73%
2006*	67%	67%
2006*	73%	60%
2006*	53%	100%
2006*	67%	80%

2006*	73%	67%
2006*	73%	80%
2006*	87%	93%
Promedio	70%	75%

* El asterisco indica que el último curso de matemáticas ha sido seguido virtualmente en la UOC

El análisis de estos datos permite enunciar las siguientes hipótesis:

- En general la realización del taller y el uso de las tablas de lectura mejoran la competencia lectora de fórmulas matemáticas, aunque hay algunos casos en los que no es así.
- Los estudiantes que hace más tiempo que no siguen un curso de matemáticas tienen un nivel más bajo de competencia lectora. En este grupo el taller y el uso de las tablas de lectura es significativamente más eficaz que en el otro grupo.
- Los estudiantes que hace menos de dos años que han realizado un curso de matemáticas, que coincide con ser un curso no presencial, tienen un porcentaje de aciertos inicial del 70% que no mejora significativamente con el uso del taller.

Es claro que este primer estudio a partir de la realización del taller necesitaría ampliarse y profundizarse pero creemos que es suficiente para detectar la necesidad de este tipo de talleres en el caso de estudiantes que hace tiempo no siguen cursos de matemáticas (presenciales o virtuales). Durante el taller, además, algunos estudiantes comentaron que la mejorar su competencia lectora les permitía sentirse más seguros en la asignatura de matemáticas y mejorar su aprendizaje virtual.

En la segunda parte de la prueba, los estudiantes de tres aulas de Análisis Matemático han utilizado un material web con las locuciones incrustadas allá donde se definía por primera vez una nueva fórmula matemática. Se les ha pasado una encuesta al final del semestre para conocer su opinión y grado de satisfacción sobre la prueba piloto de las locuciones. La encuesta consiste en valorar de 1 a 5 según la escala de Likert² las afirmaciones siguientes:

1. Considero que las locuciones de fórmulas insertadas en el material web han sido imprescindibles para aprender la forma de leer estas fórmulas.
2. Sin las locuciones de las nuevas fórmulas matemáticas no habría sabido como leerlas correctamente.
3. Aconsejaría que en todas las asignaturas de matemáticas de la universidad se pudiera tener las locuciones de las fórmulas matemáticas.

² 1 = totalmente de acuerdo, 2 = en desacuerdo, 3 = ni de acuerdo ni en desacuerdo, 4 = de acuerdo, 5 = totalmente de acuerdo.

6 M Antonia Huertas¹, Mireia Pascual³, César Córcoles¹, Laia Llorens^{P2P} y Roger Griset²

Han respondido 15 estudiantes. El Promedio de las valoraciones hechas por los 15 estudiantes en la encuesta sobre el material web con locuciones de fórmulas está en la tabla 2.

Table 2. Resultado de la encuesta sobre el material web con locuciones de fórmulas.

Afirmación	Valoración
1.	3,53
2.	3,60
3.	4,27

La información que nos proporciona es que los estudiantes que han contestado consideran que:

- Las locuciones de fórmulas han ayudado a saber/tener seguridad en la forma de leer esas fórmulas.
- Todas las asignaturas de matemáticas de la UOC se beneficiarían de tener locuciones de fórmulas matemáticas.

La experiencia de esta primera prueba piloto de incorporación de fórmulas locutadas en un material didáctico en web nos dice que hay una necesidad educativa que no estaba cubierta pero también se ha visto que la solución es parcial. Por una parte los estudiantes han mostrado satisfacción confirmando que las locuciones les han ayudado a tener un conocimiento mejor del lenguaje matemático utilizado y como consecuencia una mayor seguridad en su uso. Pero por otro lado el coste en tiempo de dedicación de cinco personas, (el experto en matemáticas, la locutora, corrección lingüística, técnico en audiovisual y en tecnología web), en relación a la cantidad final de fórmulas locutadas ha sido grande, y sin embargo muchas fórmulas necesarias han quedado aún sin "voz".

Como consecuencia nos hemos formulado la siguiente preguntas: ¿Sería interesante hacer un repositorio de contenidos abiertos de fórmulas matemáticas locutadas y multilingüe? ¿Qué retos nos plantea y cuáles serían los puntos a resolver? ¿Podría este repositorio estar en un entorno colaborativo en el que se puedan hacer aportaciones de forma distribuida? La respuesta entusiasta a esta serie de preguntas es que creemos que sí es interesante aunque haya dificultades y retos en el camino (*a quina pregunta?*). En la siguiente sección presentamos el proyecto de investigación e implementación de este repositorio abierto.

3 Repositorio abierto de locuciones multilingüe

El objetivo del proyecto es la construcción de un repositorio de locuciones de fórmulas matemáticas básicas para estudios de ingeniería y ciencias de primeros cursos de universidad. El repositorio debe ser de fácil acceso para los usuarios, independiente de un material determinado y por tanto reusable. Se pretende que sea multilingüe, aunque empezaremos con una prueba piloto con dos lenguas (castellano y catalán). También queremos que sea de acceso abierto a cualquier usuario con conexión a internet para que pueda utilizarlo, sin tener en cuenta si pertenece o no a una comunidad universitaria.

Un repositorio de fórmulas matemáticas locutadas tiene la ventaja que la notación matemática es un lenguaje universal y lo único que cambia es cómo se lee, que depende de la lengua concreta del usuario.

En la elaboración de este repositorio, y para facilitar la futura integración en materiales que sigan estándares web de matemáticas, es importante utilizar en la medida que sea posible una taxonomía compatible con la que proporciona la propia estructura del lenguaje MathML, que seguiremos como estándar web para contenido matemático. El MathML (o Mathematical Markup Language) es un lenguaje de marcado basado en XML, cuyo objetivo es expresar notación matemática de forma que distintas máquinas y aplicaciones informáticas puedan entenderla y trabajar con ella, para su uso en combinación con XHTML en páginas web, y para intercambio de información entre programas de tipo matemático en general.

La taxonomía³ de MathML se divide en tres partes: elementos de presentación o *presentation elements*, elementos de contenido o *content elements* y elementos de interfaz o *interface elements*. Los elementos de presentación describen la estructura de la notación matemática (por ejemplo el elemento *mrow* que describe una fila horizontal de caracteres, el elemento *msup* que describe un superíndice). Los elementos de contenido describen directamente los objetos matemáticos al margen de la notación concreta del objeto (por ejemplo el elemento *plus* que describe un operador con dos argumentos, el elemento *vector* que describe un vector del álgebra lineal). Los elementos de *interface* hacen referencia a requerimientos para una interfaz de navegador para MathML, y por tanto externo a la representación de la semántica o la sintaxis del lenguaje MathML.

Para la taxonomía de fórmulas y expresiones matemáticas para locutar que queremos definir a partir de la del MathML vamos a considerar solamente los elementos de contenido, que reprendan los objetos mismos. Partiremos del diccionario de contenido y operadores de MathML⁴ que hace una primera clasificación según el tema o área de las matemáticas a la que pertenece el objeto. Estos diccionarios de contenido son:

³ <http://www.w3.org/TR/REC-MathML/chapter2.html#sec2.1.1>

⁴ <http://www.w3.org/TR/MathML3/chapter4.html#contm.mathml3.operators>

- **basic_content_elements** para los elementos básicos. Contiene los símbolos *interval*, *inverse*, *lambda*, *compose*, *ident*, *domain*, *codomain*, *image*, *piecewise*, *piece*, *otherwise*.
- **algebra-logic** para aritmética, algebra y lógica. Con los símbolos *quotient*, *factorial*, *divide*, *max*, *min*, *minus*, *plus*, *power*, *rem*, *times*, *root*, *gcd*, *and*, *or*, *xor*, *not*, *implies*, *forall*, *exists*, *abs*, *conjugate*, *arg*, *real*, *imaginary*, *lcm*, *floor*, y *ceiling*.
- **relations** para relaciones (*eq*, *neq*, *gt*, *lt*, *geq*, *leq*, *equivalent*, *approx*, y *factorof*).
- **calculus_veccalc** para cálculo y cálculo vectorial (*int*, *diff*, *partialdiff*, *divergence*, *grad*, *curl*, y *laplacian*).
- **sets** para teoría de conjuntos (*set*, *list*, *union*, *intersect*, *in*, *notin*, *subset*, *prsubset*, *notsubset*, *notprsubset*, *setdiff card*, y *cartesianproduct*).
- **sequences_series** para secuencias y series (*sum*, *product*, *limit*, y *tendsto*).
- **specfun** para funciones clásicas (*exp*, *ln*, *log*, *sin*, *cos*, *tan*, *sec*, *csc*, *cot*, *sinh*, *cosh*, *tanh*, *sech*, *csch*, *coth*, *arcsin*, *arccos*, *arctan*, *arccosh*, *arccot*, *arccoth*, *arccsc*, *arccsch*, *arcsec*, *arcsech*, *arcsinh*, y *arctanh*).
- **statistics** para estadística (*mean*, *sdev*, *variance*, *median*, *mode*, *moment*, y *momentabout*).
- **linear_algebra** para álgebra lineal (*vector*, *matrix*, *matrixrow*, *determinant*, *transpose*, *selector*, *vectorproduct*, *scalarproduct*, *outerproduct*).
- **constants** para constantes y elementos de símbolos (*integers*, *reals*, *rationals*, *naturalnumbers*, *complexes*, *primes*, *exponentiale*, *imaginaryi*, *true*, *false*, *emptyset*, *pi*, *eulergamma*, *infinity*).

Algunos elementos, como son los vectores, representan objetos matemáticos, mientras que otros representan funciones u operaciones como la suma, pero no son símbolos o representaciones sino que son conceptos con significado, con semántica. Por tanto, los descriptores son términos o unidades lingüísticas que expresan conceptos y susceptibles de ser usados para definir ontologías, mapas conceptuales, taxonomías o tesauros. Aprovecharemos esta posibilidad para construir una taxonomía compatible con MatnML para nuestro repositorio.

Previa a la definición de esta taxonomía creemos en la necesidad de desarrollar una ontología que permita identificar los conceptos claves de la representación matemática con MathML y las relaciones entre ellos. Será especialmente importante la representación de la composición de los diferentes elementos para proporcionar nuevos. Como consecuencia se obtendrá una ontología de la parte de las matemáticas que este lenguaje permite expresar semánticamente. Se utilizará el software *Protege*⁵ como herramienta para la creación de la ontología y el razonador *Racer*⁶ para comprobar la consistencia de la misma.

⁵ <http://protege.stanford.edu/>

⁶ <http://www.racer-systems.com/>

Una vez definida la taxonomía y realizadas las locuciones de las fórmulas en una primera lengua, será necesario implementar interfaces de usuario que permita al estudiante acceder a la locución que precise. En la definición de su arquitectura se debe tener en cuenta que queremos permitir la ampliación a diferentes lenguas (multilingüe) y la ampliación en cantidad de locuciones en una misma lengua. El repositorio que nos interesa construir debe tener un uso cómodo, rápido y fácil, que proporcione una ayuda puntual y instantánea al usuario de cualquier material de matemáticas en la lengua de ese repositorio.

4 Conclusiones

La introducción de locuciones de fórmulas matemáticas integradas en un material web de contenido matemático para estudiantes de Ingeniería ha demostrado ser una manera fácil y eficaz de transmitir la lectura adecuada de fórmulas matemáticas y, por tanto, de facilitar el aprendizaje de su verbalización, especialmente necesario en los estudios a distancia.

Aunque esta experiencia ha sido exitosa, se han observado al menos dos cuestiones a mejorar. Por un lado el coste de locución de las fórmulas se ve aumentado por el coste en tiempo y dedicación de expertos en la inserción en un material concreto, que además exige que el material esté previamente digitalizado, y preferiblemente sea accesible mediante navegador. Por otro lado, la verbalización en diferentes lenguas, en nuestro caso catalán, castellano e inglés es necesaria, y ello independientemente de la lengua que se esté usando en el curso.

Por tanto, existe la necesidad de multilingüismo y de independizar el acceso a la verbalización de una fórmula del material de estudio concreto, ya que la notación matemática es universal, lleva a pensar en la solución de un repositorio de locuciones de fórmulas multilingües.

Es cierto que en la actualidad tecnológicamente es posible abordar el problema de la automatización de la verbalización de fórmulas, pero en el caso de las construcciones matemáticas aún hay muchas dificultades que resolver para que se pueda automatizar con seguridad. Mientras tanto, una solución factible es la construcción de un repositorio con locuciones grabadas por expertos pero respetando los estándares web de material digital de matemáticas. Esa es la propuesta que nosotros hacemos.

Referencias

1. Coll, C. (2003). Tecnologías de la información y de la comunicación y prácticas educativas. En C. Coll (Coord.). Psicología de la Educación (pp. 32-41). Barcelona: EDIUOC.
2. Mayer, R. (2005). The Cambridge handbook of Multimedia Learning. Cambridge: Cambridge University Press.

10 **M Antonia Huertas¹, Mireia Pascual³, César Córcoles¹, Laia Llorens^{2P} y Roger Griset²**

3. van Westrienen, G., Lynch, C.A. (2005). Academic Institutional Repositories: Deployment Status in 13 Nations as of Mid 2005. *D-Lib Magazine* 11(9). Available online at [doi:10.1045/september2005-westrienen](https://doi.org/10.1045/september2005-westrienen).