Trayecto para Precisar Heurísticas en un Modelo Conceptual

Graciela D.S. Hadad[1][0000-0003-4909-9702], Jorge H. Doorn[1], María Celia Elizalde[1][0009-0007-9181-9829] y Marcela N. Ridao[2][0000-0002-1281-4613]

[1] Instituto de Ingenierías y Nuevas Tecnologías, Universidad Nacional del Oeste, Argentina
[2] Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina
{ghadad,jdoorn,melizalde}@uno.edu.ar; mridao@exa.unicen.edu.ar

Resumen. Los ingenieros de requisitos deben adquirir información acerca del contexto donde se desempeñará el futuro sistema de software y registrarla en modelos apropiados. Estos modelos son utilizados durante la definición de los requisitos del sistema de software, pero también sirven de referencia para atender cualquier duda que pueda surgir en etapas posteriores del desarrollo. Desafortunadamente, se ha comprobado que estos modelos dependen muy fuertemente de quien los construya. Esto hace necesario que procedimientos y recomendaciones que acompañan a los modelos consideren factores cognitivos, de manera que los resultados no sean tan dependientes de las personas. En tal sentido, se ha reformulado un proceso de construcción de un modelo en lenguaje natural con heurísticas que atienden dichos aspectos cognitivos y semánticos. Se presenta un experimento que utiliza este proceso modificado y se lo compara con experimentos antecesores donde se usó un proceso sin consideraciones lingüísticas-cognitivas y un proceso que atiende solo algunas cuestiones lingüísticas para construir el mismo modelo. Se pudieron evidenciar mejoras en la calidad del modelo cuando se utiliza un proceso con heurísticas más precisas que atiende tanto aspectos semánticos y pragmáticos como cognitivos.

Palabras Clave: Ingeniería de Requisitos, Modelado Conceptual, Glosario, Factores Cognitivos.

1 Introducción

Se ha observado, a través de diversos estudios [1, 2, 3, 4], la existencia de defectos en los modelos utilizados en la Ingeniería de Requisitos (IR). Aunque muchos de ellos pueden identificarse mediante verificaciones y validaciones, otros permanecen latentes hasta la puesta en servicio del software que los implementa. Sin embargo, una mejora sensible podría obtenerse con un tratamiento más cuidado, tanto en la elicitación como en el modelado, incorporando guías más precisas.

El modelado produce una transformación de la información recolectada hacia la información registrada, utilizando un proceso de abstracción y elaboración, por lo que es importante cómo se manifiesta la información recolectada. Ella puede estar en notas de

Proceedings of the 28th Workshop on Requirements Engineering (WER2025),
August 20-22, 2025, Rio de Janeiro-RJ, Brazil.
https://doi.org/10.29327/1588952.28-17

los elicitadores, un resumen o acta, la grabación de una entrevista o reunión, o su transcripción, entre otras, incluso valerse de la memoria de los elicitadores. Esto hace notar la relevancia de esos registros de entrada al modelado y de su posterior procesamiento.

Se han realizado estudios sobre la calidad de un modelo conceptual en particular: el Léxico Extendido del Lenguaje (LEL) [5], habiéndose detectado tanto problemas de completitud, como de información irrelevante e inconsistente o artificiosa, ya que no tiene correlato con la realidad [6, 7]. Gran parte de estos defectos provienen de heurísticas poco precisas para la construcción del modelo, no contemplando cuestiones semánticas y pragmáticas. Esto ocurre no solo por ser un modelo escrito en lenguaje natural, sino también por cuestiones de índole cognitiva presentes en cualquier actividad desarrollada por personas [8].

En este artículo se presentan los efectos de las mejoras introducidas en el proceso de construcción del modelo LEL a través de cinco experimentos clave, seleccionados de un conjunto más amplio. A partir de un caso inicial se realizaron sucesivas propuestas de mejoras y cada una de ellas fue puesta a prueba mediante un experimento controlado. De los resultados de cada experimento se derivaron mejoras que se fueron incorporando al proceso. A partir del último experimento, que se detalla en este artículo, se elaboraron y reforzaron recomendaciones de carácter general para construir modelos conceptuales, las que pretenden atender ciertos factores cognitivos influyentes. Estos estudios se enfocaron en el modelo LEL dado su difusión en la literatura [9, 10, 11] y su aplicación en la industria [12, 13, 14].

En la sección siguiente se describen factores cognitivos que pueden afectar a los ingenieros que participan en un proceso de requisitos, el tratamiento de la información a ser modelada y el modelo LEL utilizado en los experimentos. En la sección 3 se describe el proceso reformulado de construcción del LEL a través de sus sucesivas mejoras en función de cada experimento previo, incluyendo cuestiones cognitivas. En la sección 4 se describe el último experimento, donde participan sujetos con distintos roles, se analizan sus resultados, los que se comparan con los resultados de experimentos anteriores, y se refuerzan pautas generales para modelar. Finalmente, se presentan conclusiones y futuros trabajos.

2 Marco Conceptual

Las dificultades que se afrontan al construir un modelo a partir de información adquirida en una entrevista dependen de las características propias del modelo, de los registros intermedios que se realicen durante o inmediatamente después de la entrevista y de las habilidades y disposición de quien elabora el modelo. Por ello, a continuación, se resumen los factores cognitivos involucrados, las formas de transcripción de entrevistas para utilizar registros rigurosos al modelar y se describe el modelo conceptual utilizado en el experimento.

2.1 Factores Cognitivos

La intervención de las personas con sus propias habilidades introduce cuestiones que pueden beneficiar o entorpecer las actividades y sus resultados. Por ello es necesario

también considerar los factores cognitivos intervinientes de manera de potenciar el desempeño de las personas [15].

Los factores cognitivos son los procesos mentales que permiten a las personas recibir, interpretar y elaborar la información que reciben [16]. La información debe ser percibida, atendida, recordada y organizada de manera tal que guíe apropiadamente la toma de decisión, el juicio y la acción [17].

Estos factores influyen en cómo las personas aprenden, se desempeñan y responden frente a estímulos [15], pudiendo impactar positiva o negativamente. Estos factores involucran diversas funciones cognitivas, tales como la atención, la memoria, la percepción, el razonamiento y el lenguaje, entre otras. Es importante disponer de pautas que puedan mitigar desvíos en estas funciones cognitivas.

2.2 Transcripción de Entrevistas

Es habitual el uso de entrevistas con clientes, usuarios y expertos del dominio para elicitar información sobre el contexto de aplicación y sus necesidades [4, 18, 19], para construir modelos a partir de esa información. Muchas son las razones que hacen necesario tener registros fidedignos de esas entrevistas. Claramente, se debe evitar recurrir a la memoria o a notas precarias al construir un modelo ya que esto incentiva fuertemente el uso del conocimiento previo del entrevistador proveniente de otros contextos. Además, algún detalle puede no ser percibido o ser mal interpretado y, de carecer de la fuente primaria, resultará imposible su corrección a posteriori. Eventualmente, puede ocurrir que quien modele no sea quien haya participado en la entrevista, aunque los posibles inconvenientes que puedan surgir de esta situación no han sido ponderados apropiadamente en la literatura.

Las entrevistas son uno de los abordajes clásicos en muchas de las investigaciones cualitativas de las Ciencias Sociales [20, 21, 22]. En estas disciplinas se presta una especial atención al tratamiento posterior de la entrevista, incluyendo la relación entre el entrevistador y el interpretador de la entrevista [22]. Este factor implica tener en cuenta si se trata de la misma persona o no y el lapso transcurrido entre la entrevista y su interpretación. Además, se considera mandatorio grabar y transcribir las entrevistas, al extremo que no se considera válida una afirmación que no pueda referenciarse con precisión a un punto específico de una transcripción [20, 21].

Existen diversas variantes en una transcripción [23], tales como: transcripción literal, transcripción literal reducida, transcripción editada y transcripción comentada. La *transcripción literal* registra cada detalle de la entrevista tal como ocurrió, incluyendo pausas, interjecciones, risas, toses, etc., así como indicaciones de ironía, duda, enojo y otras. La *transcripción literal reducida* elimina todos esos elementos del diálogo y eventualmente corrige errores gramaticales; se considera que este tipo de transcripción puede inducir, en algunos casos, a interpretaciones erróneas. La edición de la transcripción puede aplicarse tanto a la transcripción literal como a la transcripción literal reducida y consiste en eliminar muletillas, frases irrelevantes y otros componentes poco importantes. La *transcripción comentada* coloca agregados que faciliten la lectura de la trascripción. Esto muestra lo cuidadosa que debe ser la transcripción y aunque se usen herramientas automatizadas es importante un ajuste manual.

2.3 Modelo Léxico Extendido del Lenguaje

El modelo utilizado en los experimentos fue el Léxico Extendido del Lenguaje [5], el cual representa el vocabulario utilizado en el contexto de aplicación a través de la descripción de términos o símbolos relevantes empleados por los clientes y usuarios, y por documentación u otro material disponible. Cada símbolo se compone de uno o varios nombres (sinónimos), una o varias oraciones en el componente Noción (denotación) y una o varias oraciones en el componente Impacto (connotación). La Noción está conformada por una *definición* y una o más ampliaciones. La *definición* debe seguir la forma aristotélica, *género próximo* (concepto más general cercano al símbolo) y *diferencia específica* (característica fundamental del símbolo y que lo distingue de otros).

Para una descripción más homogénea y completa de los símbolos, estos se clasifican en cuatro tipos: Sujetos (entidades activas), Objetos (entidades pasivas), Verbos (actividades) y Estados (condiciones de los sujetos, objetos o verbos). La Tabla 1 presenta la información que se describe en cada componente del símbolo.

Tabla 1. Contenido en el símbolo según su tipo.

Tipo	Ampliaciones en la Noción	Impacto
Sujeto	Indicar rol o responsabilidad	actividades que realiza o recibe
Objeto	Indicar otros objetos con los que se relaciona	actividades que se le aplican o que se realizan con él
Verbo	Indicar quien lo ejecuta, cuando y donde se realiza	acciones, operaciones o procedimientos involucrados en la actividad, situaciones que impiden su realización, y otras actividades o situaciones que desencadena
Estado	Indicar qué estados y actividades han conducido a este estado	otros estados y actividades que pueden ocurrir a partir de este estado

3 Proceso Reformulado para la Construcción del LEL

El proceso original de construcción del modelo LEL proponía elicitar información, elaborar una lista de términos, que se clasificaban y describían; y posteriormente el LEL se verificaba y validaba [24]. Como se observa en los trabajos [5] y [24], esta lista era relativamente inmutable, ya que no existían previsiones ni heurísticas que dieran pautas acerca de cómo ampliarla; de esta manera su efecto neto era el de asegurar que todo término detectado inicialmente fuera evaluado como posible integrante del modelo, pero a su vez actuaba como un mecanismo inhibitorio para buscar nuevos términos.

Según se desprende de estos trabajos [5, 24], la construcción del modelo LEL no necesariamente estaba restringida a la información obtenida de una sola fuente de información ni se tenía en cuenta la contemporaneidad del tratamiento de las varias fuentes, si las hubiera. Sin embargo, se elaboraba una lista de inicial de términos para cada fuente de información.

Se realizó un primer experimento, aplicando esta versión del proceso de construcción del LEL [24] con un caso real basado en un documento legal que describía toda la

gestión del Plan de Ahorro y Préstamo para la adquisición de vehículos 0km. Participaron 9 grupos que construyeron, cada uno de manera independiente, un LEL a partir de una capacitación previa sobre el modelo y el proceso de construcción original, cuyos resultados se describen en [6, 1]. Debido al nivel alto de incompletitud presentado por todas las muestras del LEL, calculado aplicando el método Detection Profile Method [25], se realizó un análisis semántico, basado en el reconocimiento de sinónimos y homónimos entre muestras, previo a la estimación de completitud, detectándose apenas una leve mejoría, presentado en [26]. El mejor grupo mostraba una completitud del 51%, con 5 grupos por debajo del 30%. Asimismo, se observaron símbolos no relevantes (triviales o fuera de alcance) y nombres de símbolos inventados (no existentes en el caso) [26, 27].

Dado este diagnóstico que no solo mostraba niveles pobres de completitud sino además baja coincidencia de símbolos entre las 9 muestras del LEL, se definió un proceso más detallado, con actividades bien delimitadas por etapas (Identificar símbolos, Organizarlos y Describirlos), que incluían guías para identificar sinónimos, homónimos y una heurística simple para identificar relaciones de jerarquía [28].

Se realizó un segundo experimento utilizando el mismo caso con 3 grupos de ingenieros novatos (estudiantes de grado avanzados) [29]. Se compararon las 9 muestras anteriores y las 3 nuevas, observándose que el proceso refinado permitió emparejar los LEL de 3 grupos expertos del experimento anterior con los 3 grupos de novatos, con un nivel de completitud muy similar; esos 6 mejores grupos tuvieron un nivel de completitud entre 38% y 51%. Esto llevó a un estudio semántico más profundo sobre el tratamiento de relaciones taxonómicas entre símbolos y la ocurrencia de nominalizaciones de verbos (verbos en forma de sustantivos), lo que llevó a un nuevo refinamiento de las heurísticas de identificación y descripción de símbolos.

Se detalló con más precisión el concepto de símbolo genérico y especializado, jerarquías incompletas y verbos nominalizados, junto con el manejo de oraciones complejas para evitar ambigüedades y la coherencia en la descripción de símbolos según su tipo para reducir omisiones [30]. Ello llevó a un tercer experimento donde se construyeron 3 nuevas muestras del LEL para el caso Plan de Ahorro (logrando un total de 15 muestras) y 4 muestras del LEL para otro caso Producción de Cajas de Cartón, donde si bien hubo mejoras de completitud en las muestras nuevas, se detectaron un número inaceptable de símbolos inventados (5 de 15 muestras con 20 a 40% de símbolos inventados), persistiendo la baja coincidencia entre muestras [31].

Se planteó la necesidad de considerar las cuestiones cognitivas que pueden inducir a que los ingenieros omitan información, incorporen información inconsistente con la realidad e información irrelevante. En un primer paso se reformuló el proceso hacia un proceso iterativo para la selección de un número acotado de símbolos y su descripción en cada iteración, junto con el uso de la técnica de cortar y pegar desde la fuente textual con información elicitada [7]. Se realizó un cuarto experimento utilizando el nuevo proceso iterativo y el proceso anterior para el caso Transporte Público en Bicicleta. Se observaron menos símbolos omitidos (12,5%), menos irrelevantes (3,6%) y más símbolos correctos (96,4%), pero persistieron *sinónimos* inventados, descripciones inventadas y omitidas [7]. La Fig. 1 resume los antecedentes del proceso reformulado.

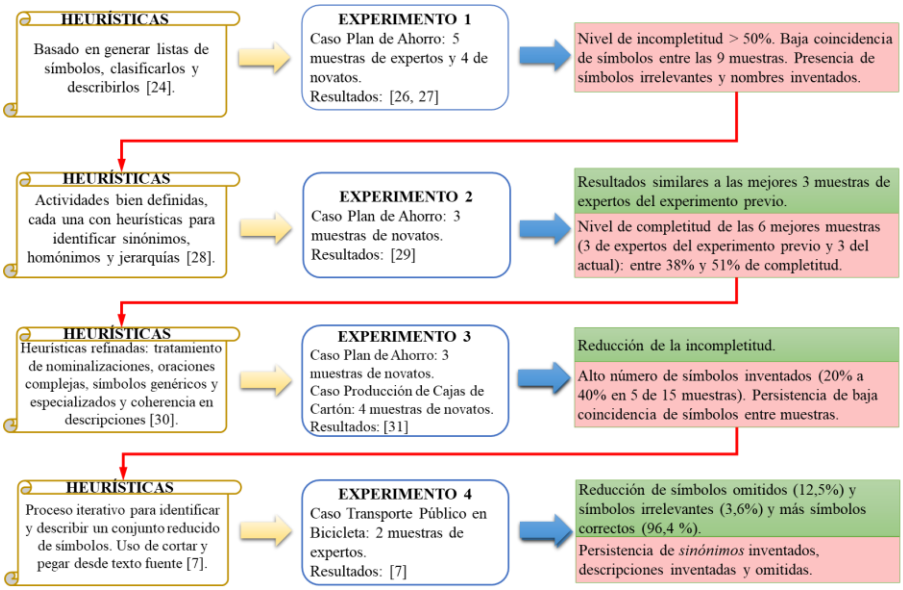

Figura 1. Resumen de Experimentos Previos.

Pese a las mejoras introducidas resultó visible la persistencia de un número significativo de defectos residuales en los modelos LEL construidos, tales como: i) omisiones de sinónimos, homónimos, estados ocultos, relaciones jerárquicas y otras, ii) ambigüedades debido a homónimos integrados en un símbolo, a símbolos genéricos o especializados no diferenciados, entre otros, e iii) inconsistencias por símbolos artificiosos (sin correlato en el contexto de aplicación). Por ello, se fueron precisando y agregando heurísticas al proceso, referidas a la selección de símbolos y su descripción, principalmente atendiendo cuestiones semánticas y pragmáticas. Estas heurísticas involucraron los siguientes aspectos[1]:

- Registrar en forma precisa las relaciones taxonómicas y mereológicas (todo-parte) existentes entre los símbolos. Esto se realizó utilizando frases específicas, tales como "es parte de", "es un", "es el" o "puede ser", para indicar implícitamente la posición de cada símbolo en la relación correspondiente.
- Registrar en forma precisa los homónimos. Esto se realizó cuidando qué información corresponde a cada símbolo homónimo, tanto en la noción como en el impacto.
- Registrar en forma precisa las nominalizaciones respecto a verbos cuando se usa su forma nominal, identificando la existencia de sinónimos en su forma verbal o de homónimos (tipo objeto o sujeto) en su forma nominal.
- Registrar en forma precisa los homónimos generados por el uso de apócopes de especializaciones que se confunden con su genérico. Esto se realizó indicando qué sinónimos posee cada homónimo.

[1] La descripción del proceso y sus heurísticas junto con ejemplos se encuentran en: https://drive.google.com/file/d/1f7GFhkjkTR_XRDMyn3z5r0TI27KV7lwo/view?usp=sharing

- Registrar en forma precisa las especializaciones anónimas sustituidas por el genérico. Esto también se realizó indicando qué sinónimos posee cada homónimo.
- Registrar en forma precisa las diferencias entre los estados y las relaciones de especialización permanentes. Esto se realizó asegurando que los estados estén desacoplados en su descripción de los sujetos, objetos o verbos relacionados.

Finalmente, el proceso se reformuló integralmente, teniendo como base iteraciones acotadas donde se seleccionan unos pocos términos (5 o 6), se los clasifica y describe, y luego se pasa a seleccionar otros pocos para describirlos y así sucesivamente hasta cumplir con el criterio de parada establecido, dando por terminado el registro de la información recabada de esta fuente de información. El objetivo principal fue mitigar los factores cognitivos presentes en esta actividad de modelado. Las características esenciales del proceso reformulado son:

- Identificación y selección de grupos acotados de términos a modelar en iteraciones, para centrar la atención en una cantidad razonable de elementos.
- Reglas de modelado para los símbolos y sus relaciones (sinónimos, homónimos, relaciones taxonómicas, relaciones mereológicas, vocabulario restringido), que resultan necesarias dada la variedad en las expresiones del lenguaje natural. Estas reglas son las referidas a las cuestiones semántico-pragmáticas antes indicadas.
- Uso de la técnica de copiar y pegar desde la información elicitada al modelo, evitando malinterpretaciones del ingeniero de requisitos en la traducción al modelo o interpretaciones basadas en su conocimiento previo.
- Criterio para el manejo del fin de una iteración en el modelado, para limitar el trabajo a partes comprensibles.
- Criterio de parada del proceso, para reducir la incertidumbre al establecer el nivel de calidad alcanzado del modelo.

4 Trabajo Realizado y Resultados

Dado los resultados de los experimentos antecesores, se diseñó un nuevo experimento tomando un caso real, donde participaron tanto sujetos con experiencia en el ámbito informático, particularmente en Ingeniería de Requisitos, como novatos. Se describe el experimento y sus resultados, los que se comparan con resultados de experimentos previos, y se argumenta sobre la validez del experimento. Asimismo, se presentan recomendaciones generales para el modelado considerando factores cognitivos, con base en los resultados obtenidos.

4.1 Experimento con Expertos y Novatos

El experimento fue diseñado para estudiar si se lograban mejoras en la calidad del modelo LEL con el proceso reformulado, utilizando el caso Gestión de una Clínica Odontológica. Para que el caso sea lo más real posible, se realizó previamente una entrevista abierta a un usuario en su lugar de trabajo, con el fin de que este describiera las actividades que realiza y sus necesidades, la que fue grabada y transcripta. Esta información

elicitada fue la entrada para modelar un LEL, donde participaron 7 sujetos, siendo 3 de ellos novatos (estudiantes de grado). Los sujetos expertos (profesionales y académicos) tenían experiencia en realizar entrevistas y modelar el LEL con el proceso original, y fueron capacitados en el nuevo proceso de construcción del LEL. A los sujetos novatos se los entrenó en el modelo LEL y en el proceso de construcción, mediante clases explicativas, ejemplos y consultas. A todos los sujetos se les entregó un documento con la descripción del modelo LEL y el nuevo proceso de construcción, que incluía ejemplos de aplicación de las heurísticas. Cada sujeto construyó su versión del modelo LEL de manera independiente de los otros, aplicando la plantilla del modelo y completando una planilla para el registro de tiempos.

Todos los sujetos fueron informados del propósito del experimento y prestaron su acuerdo en forma explícita. Los datos personales y de la institución participante fueron almacenados en condiciones seguras y su anonimato garantizado. No hubo ninguna actividad que involucrara personal vulnerable ni existió ningún conflicto de intereses.

Se utilizaron dos alternativas como entrada para modelar: el audio de la entrevista en un formato estándar o la *trascripción literal* de ella. Para evitar sesgos en la transcripción, se realizó la desgrabación mediante una herramienta de reconocimiento de voz; y luego el texto fue revisado por un ingeniero, usando una herramienta de desarrollo propio, para corregir errores de desgrabación, identificar interlocutores y marcar gestos tales como pausas, interrupciones y cambios en el tono de voz.

Como se menciona en [22] respecto al factor relación entrevistador - interpretador en las Ciencias Sociales, esto mismo puede ocurrir en un proceso de requisitos entre el entrevistador y el modelador, si se trata o no de la misma persona. Por ello en el experimento se consideró la combinación de estos roles, es decir, tener sujetos que entrevistan y luego modelan, y otros que solo modelan. Además, está la variante para estos roles de qué información de entrada utilizaron: el audio de la entrevista o su texto transcripto. Entonces, en el experimento participaron los siguientes roles:

- Sujeto 1: entrevistador y modelador, usando texto, experto
- Sujeto 2: entrevistador y modelador, usando audio, experto
- Sujeto 3: solo modelador, usando audio, experto
- Sujeto 4: solo modelador, usando texto, experto
- Sujeto 5: revisor de transcripción y modelador, usando texto, novato
- Sujeto 6: solo modelador, usando texto, novato
- Sujeto 7: solo modelador, usando texto, novato

A partir del experimento se obtuvieron los datos que se muestran en la Tabla 2, donde son considerados defectos los símbolos *irrelevantes* referidos a conceptos no esenciales en el contexto de aplicación o fuera del alcance del sistema, los símbolos *inventados* o artificiosos que corresponden a aquellos no mencionados en la entrevista, y los símbolos *omitidos* que son aquellos conceptos importantes en el contexto de aplicación ausentes en el modelo. Se usó la palabra *incorrecto* para catalogar a símbolos y oraciones irrelevantes o inventadas. La evaluación de las 7 versiones fue realizada por los tres primeros autores del artículo en dos etapas: 1) evaluación independiente, generando cada uno una planilla con el detalle de los defectos detectados en cada muestra, y 2) reuniones para consensuar resultados de cada muestra.

Tabla 2. Datos de las versiones construidas del modelo por sujetos expertos y novatos.

Roles cumplidos:	Entrevistador y Modelador (Experto)		Solo Modelador (Experto)		Revisor y Modelador (Novato)	Solo Modelador (Novato)	
	Sujeto 1	**Sujeto 2**	**Sujeto 3**	**Sujeto 4**	**Sujeto 5**	**Sujeto 6**	**Sujeto 7**
Medio de Entrada:	Texto	Audio	Audio	Texto	Texto	Texto	Texto
símbolos correctos detectados	**26**						
símbolos modelados	18	25	61	45	15	21	18
símbolos modelados correctos	16	20	21	23	13	16	15
símbolos modelados irrelevantes	1	3	16	16	2	5	3
símbolos modelados inventados	1	2	24	6	0	0	0
símbolos omitidos	10	6	5	3	13	10	11
% símbolos correctos	88,9%	80,0%	34,4%	51,1%	86,7%	76,2%	83,3%
% símbolos irrelevantes	5,6%	12,0%	26,2%	35,6%	13,3%	23,8%	16,7%
% símbolos inventados	5,6%	8,0%	39,3%	13,3%	0,0%	0,0%	0,0%
% símbolos omitidos	38,5%	23,1%	19,2%	11,5%	50,0%	38,5%	42,3%
oraciones modeladas	77	115	216	263	78	78	104
oraciones modeladas correctas	75	109	130	166	70	73	94
% oraciones incorrectas	2,6%	5,2%	39,8%	36,9%	10,3%	6,4%	9,6%
Tiempo promedio de modelado por símbolo	30 min	36 min	35 min	34 min	17 min	24 min	17 min

Surge de los datos obtenidos que hubo pocos símbolos inventados e irrelevantes y pocas oraciones incorrectas en nociones e impactos, exceptuando los sujetos que solo modelaron y son expertos. En el caso de estos dos sujetos (Sujeto 3 y Sujeto 4) modelaron muchos más símbolos y oraciones que el resto, lo que los llevó obviamente a menos símbolos omitidos, pero más información irrelevante intentando completar con toda la información disponible de la fuente. Es posible que estos expertos fueran afectados por el sesgo cognitivo denominado *efecto de estereotipo* [32], en el que el experto siente el temor de no superar a los novicios perjudicando de esa manera su status. Los sujetos afectados por este sesgo pueden reducir su desempeño u objetividad. Además, el Sujeto 3 tenía experiencia en el dominio, lo que podría indicar el número alto de símbolos y oraciones inventadas por su conocimiento previo (*sesgo de confirmación*).

Notoriamente, los novatos no presentan símbolos inventados; ello podría deberse a escasez de conocimiento previo como también apegarse al proceso al detalle. Sin embargo, estos sujetos novatos son los que tienen más símbolos omitidos. No se observaron diferencias significativas entre el novato que revisó la transcripción respecto de los otros novatos que solo modelaron.

Respecto a los expertos que cumplieron el mismo rol (entrevistar y modelar, o solo modelar), quienes usaron el audio de la entrevista modelaron un porcentaje menor de símbolos correctos y lo mismo ocurrió con las oraciones correctas.

Los expertos utilizaron similar tiempo promedio de modelado por símbolo, con un leve incremento para los que usaron audio, mientras que los novatos insumieron menos tiempo. Surge como posible el *sesgo de la profecía negativa autocumplida*, en el que el novato puede pensar que no tiene ninguna oportunidad de tener un desempeño decoroso, comparado con el de los expertos [33].

4.2 Comparación con Estudios Previos

Independientemente de la experiencia de cada sujeto, disminuyeron significativamente

las omisiones respecto a estudios previos donde las omisiones fueron ampliamente superiores al 50% en todas las muestras del modelo [27], utilizando el proceso original de construcción del LEL. Con las mejoras iniciales se había logrado emparejar el nivel de completitud entre sujetos expertos y novatos [29, 31] pero sin una mejora significativa en el número de omisiones.

En la Fig. 2, se analiza la frecuencia de detección de símbolos correctos en las 7 muestras para el caso Clínica Odontológica y en las 9 muestras iniciales para el caso Plan de Ahorro. Puede observarse la baja cantidad de símbolos comunes detectados en el caso Plan de Ahorro con solo un 10% de los símbolos detectados en las 9 muestras, mientras que en el caso Clínica Odontológica, casi el 27% de los símbolos fue detectado en las 7 muestras. Por otra parte, en el caso Clínica Odontológica, una amplia cantidad de símbolos fue detectada por un buen número de sujetos: el 53% de los símbolos detectados por 5 a 7 sujetos (ver las tres columnas iniciales del gráfico izquierdo). Por el contrario, en el caso Plan de Ahorro, un gran porcentaje de símbolos fueron detectados por unos pocos sujetos (ver las últimas columnas del gráfico derecho).

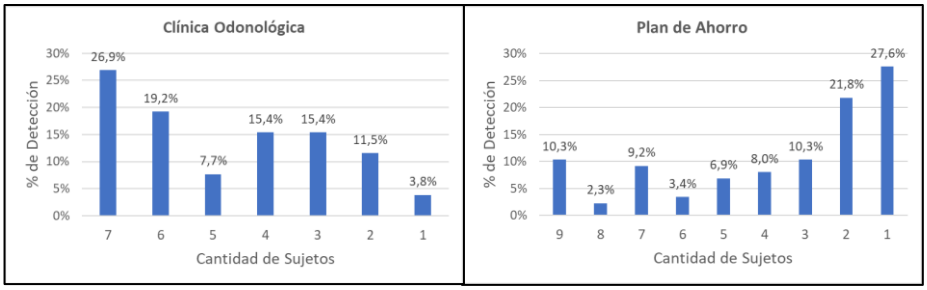

Figura 2. Frecuencia de detección de símbolos.

Comparando con experimentos antecedentes que utilizaron el proceso original del LEL aplicado al Plan de Ahorro con 15 muestras y el proceso del LEL con algunos refinamientos aplicados al caso Producción de Cajas con 4 muestras [31], se observó que en el caso Clínica Odontológica disminuyó significativamente la cantidad de muestras con símbolos inventados. En el caso Plan de Ahorro solo el 33% de las muestras (5 de 15) tuvo de 0 a 10% de símbolos inventados y en el caso Producción de Cajas el 50% (2 de 4), mientras que en el caso Clínica Odontológica, el 71% de las muestras (5 de 7) tuvo entre 0 a 10% de símbolos inventados.

4.3 Validez del Experimento

La selección de sujetos fue realizada de manera tal de contar con sujetos expertos (profesionales de informática y académicos) y sujetos novatos (estudiantes de grado de informática). Todos los sujetos habían construido al menos un LEL con anterioridad, pero ninguno había utilizado el proceso por iteraciones.

Para la capacitación, los sujetos expertos recibieron el documento del proceso para su lectura, con un período de consultas individuales sobre el proceso. Los sujetos novatos tuvieron una clase expositiva sobre el modelo LEL y el proceso reformulado, además de recibir el documento del proceso y el período de consultas individuales. La

asignación a expertos del audio o texto transcripto fue aleatoria, dado que todos ellos tenían similar formación y experiencia. Esto posibilitó tener combinaciones de roles.

Los sujetos trabajaron de manera independiente, sin ninguna comunicación con los restantes, contando con los mismos instrumentos: el documento con la definición del proceso y ejemplos, la plantilla del modelo LEL y la planilla para registrar tiempos.

Hubo variedad de roles para estudiar la influencia por participar en actividades previas vinculadas al caso: dos sujetos que habían participado en la entrevista y un sujeto que revisó y corrigió la transcripción de la entrevista, frente a los restantes que solo modelaron. Además, se combinó el medio de soporte de la información de entrada al modelado (el audio o el texto transcripto), dado que cualquiera de ambas posibilidades consideradas es factible de ocurrir en el ámbito profesional.

Los resultados obtenidos refieren a una muestra de sujetos con distinta experiencia, realizando el mismo trabajo partiendo de distinto conocimiento (participación en la entrevista, en la revisión de la transcripción o sin ningún acercamiento previo al caso) y utilizando registros diferentes de la información elicitada (audio o texto). Ello implica que los resultados sobre esta muestra variada pero pequeña solo pueden permitir esbozar la influencia de estos factores de variabilidad sobre la calidad del LEL.

4.4 Discusión de los Resultados y Recomendaciones

Del análisis de los resultados del último experimento respecto a los anteriores, se nota una mejora. No obstante, persisten algunas debilidades, lo que amerita reforzar algunas pautas de modelado e incorporar algunas otras.

Se observa una mejora cuando se usa texto transcripto frente al audio, facilitando el traslado de información elicitada al modelo. Ello además facilita preservar el vocabulario de los clientes y usuarios, evitando el riesgo de invenciones basadas en experiencias previas. Asimismo, si bien la técnica de copiar y pegar del texto al modelo es efectiva, se observó que hubo oraciones en el LEL que presentaron problemas de redacción o no se ajustaban a la sintaxis o al contenido correspondiente. Es decir, existen casos donde se requiere realizar algunas abstracciones y ajustes en la sintaxis del texto fuente para adecuarse al modelo. Por ello se necesita precisar el uso de copiar y pegar, en cuanto a en qué situaciones y cómo realizar elaboraciones sobre el texto copiado. No hacer abstracciones condujo en algunos casos a información irrelevante en el modelo o con un nivel de detalle innecesario que entorpecía la comprensión del símbolo. También se detectó que no tener presente o claro el objetivo del sistema condujo a incorporar información innecesaria o fuera del alcance.

Cabe notar que la transcripción de texto se realiza con una herramienta automatizada en un tiempo despreciable, mientras que el tiempo de revisión del texto depende de la calidad de la herramienta y la habilidad del personal junior. Asimismo, el copiar y pegar tiende a facilitar la escritura en el modelo, aun cuando haya que adaptar la redacción.

Entonces, se han delineado pautas que están referidas a modelos escritos en lenguaje natural, considerando que se dispone de texto fuente de la elicitación, teniendo en cuenta aquellos factores cognitivos más influyentes en el procesamiento de información: la atención, la memoria, el lenguaje y el razonamiento.

- Prestar atención, concentrarse en la tarea. Destinarle el tiempo y espacio apropiado,

evitando interrupciones.

- Evitar basarse en la memoria, especialmente cuando se ha participado en la entrevista o se ha trabajado en la revisión de la transcripción.
- No introducir términos propios en el modelo, basados en conocimiento previo o interpretaciones del texto. Preservar el uso de términos provenientes de la fuente.
- Procesar la información obtenida de manera lo más objetiva posible, maximizando el copiar y pegar al modelo, ajustando solo lo necesario, y solo elaborar información a partir del texto fuente que la avala. Ajustes y abstracciones al texto copiado son posibles en situaciones preestablecidas.
- Mantener rastros del modelo a la fuente y viceversa. Para ello marcar el texto fuente a medida que se usa en el modelo, pero evitando que ese texto marcado quede oculto, dado que puede contener información necesaria en otros componentes del modelo.
- Tener presente el objetivo establecido para evitar incluir información no relevante y focalizarse en lo necesario.
- Evitar apartarse de las guías de modelado (dependientes del modelo). Tenerlas a mano para repasarlas, de ser necesario.

5 Conclusiones

Se ha presentado una secuencia de mejoras en un proceso de construcción de un modelo conceptual en lenguaje natural, el LEL, que partiendo de heurísticas elaboradas en 1997 ha arribado en la actualidad a heurísticas que atienden aspectos no considerados inicialmente, incorporando un abordaje de la complejidad cognitiva de las personas intervinientes. De manera incremental, a través de sucesivos experimentos, se fueron agregando guías y precisando existentes, logrando mejoras significativas en cada paso.

Otros autores también han propuesto heurísticas para este mismo modelo [9, 10, 11, 34], las que están principalmente enfocadas en la sintaxis y semántica del modelo, pero que no presentan una evaluación de dichas guías para establecer si ellas han mejorado o no la calidad del LEL. Un ejemplo de esta situación se evidencia en el trabajo de Ruidías et al. [9], donde se muestra la construcción de un modelo LEL basado en ontologías, apoyado por una herramienta que permite detectar inconsistencias y omisiones internas con base en reglas sintácticas, y se presenta la aplicación de la propuesta mediante un caso. En [11] también se presentan seis guías de construcción del LEL con base sintáctica y algún aspecto semántico, junto con once reglas de extracción de conocimiento del LEL, que fueron evaluadas sobre su usabilidad mediante encuestas realizadas a estudiantes con experiencia laboral. En el trabajo de Roldán y Antonelli [10] se describe una gramática extendida BNF para el LEL, que fue validada por profesionales usando una herramienta colaborativa.

Las nuevas recomendaciones presentadas son de carácter general, de manera que son aplicables a diversos modelos escritos en lenguaje natural. Algunas de ellas son simples, y hasta pueden considerarse intuitivas, sin embargo, surgen de observar debilidades en la puesta en práctica, por profesionales con experiencia en el modelo, de un proceso con cuidados importantes en sus guías.

Se ha iniciado un estudio de los sesgos cognitivos que afectan a los ingenieros de

software en sus actividades [35]. Se analizará cuáles los afecta en la elicitación y el modelado, de manera de poder establecer si las recomendaciones elaboradas coinciden con el conocimiento acumulado en la psicología experimental, en el sentido de ser la forma recomendada de atenuar efectivamente esos sesgos cognitivos. De ser necesario, se elaborarán nuevas recomendaciones o se ajustarán las existentes.

Referencias

1. Ridao, M., Doorn, J.H.: Estimación de Completitud en Modelos de Requisitos Basados en Lenguaje Natural. En: Proceedings of IX Workshop on Requirements Engineering (WER06), Río de Janeiro, pp. 146–152 (2006)
2. Hadad, G.D.S., Litvak, C., Doorn, J.H., Ridao, M.N.: Dealing with Completeness in Requirements Engineering. En: Mehdi Khosrow-Pour (ed.) Encyclopedia of Information Science and Technology, 3rd edition. IGI Global, Information Science Reference, Hershey, EE.UU., pp. 2854-2863 (2015) DOI: 10.4018/978-1-4666-5888-2.ch279
3. Martínez, S.N., Oliveros, A., Zuñiga, J.A., Corbo, S., Forradelas, P.: Aprendizaje de la elicitación y especificación de requerimientos. En: Anales de XX Congreso Argentino de Ciencias de la Computación (CACIC) (2014)
4. Bano, M., Zowghi, D., Ferrari, A., Spoletini, P., Donati, B.: Teaching requirements elicitation interviews: an empirical study of learning from mistakes. Requirements Engineering Journal, 24(3), 259–289 (2019) https://doi.org/10.1007/s00766-019-00313-0
5. Leite, J.C, Franco, A.: A Strategy for Conceptual Model Acquisition. En: IEEE 1st Intl Symposium on Requirements Engineering, IEEE Computer Society Press, pp. 243-246 (1993)
6. Doorn, J.H., Ridao, M.N.: Completitud de Glosarios: Un estudio experimental. En: Proceedings of VI Workshop on Requirements Engineering (WER03), pp. 317-328 (2003)
7. Elizalde, M.C., Hadad, G.D.S., Doorn, J.H.: Incorporación de heurísticas lingüístico-cognitivas en el Proceso de Requisitos. En: Memorias de 9° Congreso Nacional de Ingeniería Informática / Sistemas de Información (CONAIISI), Mendoza, pp. 312-320 (2021)
8. Hutton, C.: Semantics and the Etymological Fallacy. Language Sciences 20(2), 189-200 (1998)
9. Ruidías, H.J., Caliusco, M.L, Galli, M.R.: Construcción basada en ontologías del Léxico Extendido del Lenguaje. En: Anales de LIII Jornadas Argentinas de Informática e Investigación Operativa (43 JAIIO) - XV Simposio Argentino de Ingeniería de Software, Buenos Aires, pp. 188-202 (2014)
10. Roldán Valdiviezo, P.M., Antonelli, L.: Definición de una Gramática para el Léxico Extendido del Lenguaje. En: Proceedings of 27th Workshop on Requirements Engineering (WER24), Buenos Aires (2024) DOI 10.29327/1407529.27-3
11. Antonelli, L, Lezoche, M., Delle Ville J.: Knowledge Extraction from the Language Extended Lexicon Glossary Using Natural Language Processing. TecnoLógicas, 27(59) (2024)
12. Mighetti, J.P., Hadad, G.D.S.: A Requirements Engineering Process Adapted to Global Software Development. CLEI Electronic Journal, 19(3) (2016) DOI 10.19153/cleiej.19.3.7
13. Mira, N.C., Boggio, M.A., Perez, S.B., Salamon A.G.: Una propuesta de análisis de problemas utilizando LEL y Escenarios en entrenamientos de equipos que gestionan situaciones de crisis. En: Memorias del 6to Congreso Nacional de Ingeniería Informática / Sistemas de Información, Mar del Plata, pp. 1041-1048 (2018)
14. Andrianjaka, R.M., Razafimahatratra, H., Mahatody, T., Ilie, M., Ilie, S., Raft R.N.: Automatic generation of software components of the Praxeme methodology from ReLEL. En: Proceedings of 24th International Conference on System Theory, Control and Computing, Sinaia, Romania, pp. 843-849 (2020) DOI 10.1109/ICSTCC50638.2020.9259737

15. Roy, E.: Cognitive Factors. En: Gellman, M.D. (eds) Encyclopedia of Behavioral Medicine. Springer, Cham (2020) https://doi.org/10.1007/978-3-030-39903-0_1116
16. Waisburd Jinich, G.: Pensamiento creativo e innovación. Revista Digital Universitaria, Universidad Nacional Autónoma de México, México, 10(12) (2009)
17. Francis, G., Rash, C.E.: Cognitive Factors. Semantic Scholar, Corpus ID: 18316850, cap. 15, pp. 619-673 (2009)
18. Oliveros, A., Antonelli, L.: Técnicas de elicitación de requerimientos. En: Anales de XXI Congreso Argentino de Ciencias de la Computación (CACIC), Junín (2015)
19. Ferrari, A., Spoletini, P., Gnesi, S.: Ambiguity and tacit knowledge in requirements elicitation interviews. Requirements Engineering Journal, Springer, 21(3), 333-353 (2016)
20. Mann, S.: The Research Interview Reflective Practice and Reflexivity in Research Processes. Palgrave Macmillan, Basingstoke, pp. 199-211 (2016)
21. Buriro, A.G., Awan, J., Lanjwani, A.: Interview: A Research Instrument for Social Science Researchers. International Journal of Social Sciences, Humanities and Education,1,1-14 (2017)
22. Edwards, R., Holland, J.: What is qualitative interviewing? Bloomsbury Academic (2013)
23. Edwards, J.A.: Principles and Contrasting Systems of Discourse Transcription. En: Edwards, J., A., Lampert, M., D., (eds.), Talking data: transcription and coding in discourse research, Taylor and Francis, Nueva York, pp 3-31 (1993)
24. Hadad, G.D.S., Kaplan, G.N., Oliveros, A., Leite, J.C.S.P.: Construcción de Escenarios a partir del Léxico Extendido del Lenguaje. En: Anales de Jornadas Argentinas de Informática (26 JAIIO), SoST'97 Simposio en Tecnología de Software, Buenos Aires, pp. 65-77 (1997)
25. Wohlin, C., Runeson, P.: Defect content estimations from Review Data. En: Proceedings of 20th International Conference on Software Engineering, Japón, pp. 400-409 (1998)
26. Litvak C.S., Hadad G.D.S., Doorn J.H.: Un abordaje al problema de completitud en requisitos de software. En: Anales de XVIII Congreso Argentino de Ciencias de la Computación (CACIC), Bahía Blanca, pp. 827-836 (2012)
27. Litvak, C.S., Hadad, G.D.S., Doorn, J.H.: Correcciones semánticas en métodos de estimación de completitud de modelos en lenguaje natural. En: Proceedings of 16th Workshop on Requirements Engineering (WER13), Montevideo, pp.105-117 (2013)
28. Hadad, G.D.S.: Uso de Escenarios en la Derivación de Software. Tesis Doctoral, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Arg. (2008) DOI 10.35537/10915/2456
29. Litvak, C.S., Hadad, G.D.S., Doorn, J.H.: Mejoras semánticas para estimar la completitud de modelos en lenguaje natural. En: Memorias del 1er Congreso Nacional en Ingeniería Informática / Sistemas de Información (CoNaIISI), Córdoba (2013)
30. Litvak, C.S., Hadad, G.D.S., Doorn, J.H.: Heurísticas para el modelado de requisitos escritos en lenguaje natural. En: Anales de XX Congreso Argentino de Ciencias de la Computación (CACIC), Buenos Aires, pp. 682-691 (2014)
31. Doorn, J.H., Hadad, G.D.S., Elizalde, M.C., García, A.R.G., Carnero, L.O.: Críticas Cognitivas a Heurísticas Orientadas a Modelos. En: Proceedings of 22nd Workshop on Requirements Engineering (WER19), Recife (2019) DOI 10.29327/1298731.22-7
32. Spencer, S.J., Logel, C., Davies, P.G.: Stereotype Threat. Annual Review of Psychology, 67, 415–37 (2016) https://doi.org/10.1146/annurev-psych-073115-103235
33. Farmer, R.E.A.: The Macroeconomics of Self-fulfilling Prophecies, 2da. ed., MIT Press (1999)
34. Antonelli, L, Rossi, G., Leite, J.C.S.P., Oliveros A.: Buenas Prácticas en la Especificación del Dominio de una Aplicación. En: Proceedings of 16th Workshop on Requirements Engineering (WER13), Montevideo, pp. 319-332 (2013)
35. Mohanani, R., Salman, I., Turhan, B., Rodríguez, P., Ralph, P.: Cognitive biases in software engineering: a systematic mapping study. IEEE Transactions on Software Engineering, 46(12), 1318-1319 (2018) DOI 10.1109/TSE.2018.2877759