

A intenção de uso da Linguagem Natural na especificação de requisitos:

um estudo comparativo entre a Argentina e o Brasil

Angélica T. S. Calazans¹, Roberto A. Paldês¹, Aldegol O. Paulino¹, Fabrício R. Freire¹, Jana P. Fraga¹, Graciela D. S. Hadad^{2,3}, Ari Melo Mariano⁴

¹Centro Universitário de Brasília, Instituto CEUB de Pesquisa e Desenvolvimento, Brasília, Brasil
{angelica.calazans, roberto.paldes}@uniceub.br

²Universidad Nacional del Oeste, Escuela de Informática, Buenos Aires, Argentina
³Universidad Nacional de La Matanza, DIIT, Buenos Aires, Argentina
ghadad@uno.edu.ar

⁴Universidade de Brasília, EPR/PPCA, Brasília, Brasil
arimariano@unb.br

Resumo — O objetivo deste trabalho é realizar um estudo comparativo sobre os fatores de influência do uso da linguagem natural na especificação de requisitos na Argentina e o Brasil. A metodologia adotada é um estudo descritivo, de abordagem quantitativa por meio de equações estruturais. O modelo teórico proposto baseia-se na Teoria Unificada de Aceitação e Uso da Tecnologia. Foi elaborado um questionário com 17 perguntas para 45 analistas de requisitos da Argentina e 50 analistas de requisitos do Brasil. O instrumento de pesquisa foi validado e pode ser reutilizado. Identificou-se que os motivos que levam os analistas a utilizar a linguagem natural na especificação é a expectativa de um rendimento elevado e a percepção que o uso da linguagem natural leva a uma maior adesão às necessidades do cliente. O estudo destaca outras expectativas podem ser confirmadas com fatores de influência para a Argentina, mas não para o Brasil.

Palavras Chave— Engenharia de Requisitos, Especificação de Requisitos, Linguagem Natural, Modelo de percepção.

I. INTRODUÇÃO

A necessidade de se obter um software com requisitos completos, consistentes, relevantes e que possam atender às necessidades do cliente [1] deixa clara a importância da Engenharia de Requisitos no processo de desenvolvimento de sistemas de software. As atividades de extração e manutenção dos requisitos, partes integrantes de um processo de Engenharia de Requisitos, criam um vínculo entre os stakeholders cliente e equipe de desenvolvimento onde as visões distintas convergem para uma linguagem comum.

Ao mesmo tempo, é primordial uma especificação de requisitos clara, o mais completa possível e sem ambiguidades para geração de documentos consistentes e de fácil entendimento por todas as partes envolvidas na construção de uma solução de software. Essas características vêm ao

encontro à proposta de utilização de Linguagem Natural (LN), pois ela possibilita a estruturação das ideias que dão base a um artefato sólido, de onde se desdobrarão as demais atividades de desenvolvimento do software.

A LN é amplamente utilizada na redação de requisitos dada a sua capacidade de descrever aquilo que é essencial, ou seja, escrever os requisitos de usuário em uma linguagem que os não-especialistas possam compreender [2]. Apesar das desvantagens de sua utilização [3] [4] [5], e até mesmo do desconhecimento de profissionais da área a respeito de condições facilitadoras do seu uso de forma mais efetivo, a LN tem a sua importância reconhecida [6].

Para avaliar o impacto da escrita em LN nos dias atuais é necessário obter a percepção das pessoas que fazem uso desse recurso, os engenheiros ou analistas de requisitos e analistas de sistemas com essa atribuição. Modelos de percepção que envolvem tecnologia têm sido estudados para diferentes áreas [7] [8] [9] através da escala de medida UTAUT (*Unified Theory of Acceptance and Use of Technology* / Teoria Unificada de Aceitação e Uso da Tecnologia). Porém os estudos sobre UTAUT estão aplicados a Tecnologia de uma maneira geral, sendo necessário adaptar este estudo ao da LN. Entender os motivos da sua adoção ou não na engenharia de requisitos é indispensável a melhoria da técnica, assim como da relação das soluções trazidas pelos engenheiros e os usuários envolvidos nas diferentes etapas do processo.

No Brasil, a escala de medida UTAUT já foi utilizada para avaliar o impacto da escrita em LN nos dias atuais [10] e seu uso em diferentes contextos culturais permite uma aproximação acerca do comportamento de ambos os países em relação a LN. Em um mundo onde a interação entre profissionais ocorre de modo global, entretanto, há a necessidade de uma análise mais ampla, envolvendo

profissionais de culturas diversas, onde os aspectos semelhantes e divergentes nos possíveis stakeholders de países distintos sejam compreendidos. Isso possibilita o questionamento o quão diferente podem ser as percepções de profissionais da Argentina e do Brasil que atuam na área de requisitos a respeito da utilização da LN na construção de uma especificação de requisitos de software.

Considerando o exposto, a presente pesquisa objetiva responder as seguintes questões: quais são as percepções dos profissionais dos países analisados acerca da LN na especificação dos requisitos de um produto de software? Essas percepções variam, considerando as visões dos profissionais da Argentina e do Brasil? Para responder a essas questões, o objetivo geral desse trabalho é analisar e comparar a influência das expectativas de desempenho e esforço, bem como condições facilitadoras, com relação a utilização da LN na especificação de requisitos de dois países. É interessante ressaltar que na pesquisa não foram analisadas técnicas específicas de representação em linguagem natural.

No intuito de estabelecer essa comparação, este trabalho foi elaborado aplicando os mesmos critérios já utilizados no estudo realizado no Brasil [10]. Para isso, na seção 2 são descritos, de forma breve, os conceitos relacionados a LN, na seção 3 a metodologia da pesquisa. Na seção 4 os resultados obtidos de cada país, sua análise e comparação. Finalmente, na seção 5 são realizadas as considerações finais e conclusões.

II. OS REQUISITOS E A LINGUAGEM NATURAL

Os requisitos podem ser definidos como as condições do que o software deve cumprir: os serviços que ele deve oferecer e as restrições ao seu funcionamento [3]. O processo de Engenharia de Requisitos é composto de várias etapas, entre elas descobrir (elicitar), analisar, documentar os requisitos e verificar as funções e restrições definidas [11].

A elicitação de requisitos engloba a identificação, captura e integração de requisitos decorrentes da comunicação entre um grupo de analistas e as partes interessadas, gerando descrições textuais e/ou gráficas que refletem os conceitos mais relevantes do domínio para o desenvolvimento de uma aplicação [12].

Para facilitar essa comunicação, os requisitos são geralmente escritos e mantidos em LN, que é uma forma de comunicação compartilhada por todos os atores do processo, ainda que a flexibilidade da LN traga também riscos de ambiguidade e desentendimento.

No contexto da presente pesquisa, considera-se a Linguagem Natural o uso de frases bem estruturadas para descrever os requisitos de um sistema. Uma LN é elaborada naturalmente pelo ser humano, sem rigores de preparação [13]. É o idioma normalmente utilizado na comunicação em uma comunidade. O padrão IEEE 29148 [14] aponta diversas propriedades para que a especificação de requisitos de software obtenha um bom nível de qualidade entre elas a isenção de ambiguidade e o uso da LN para a descrição de requisitos.

Existem vários motivos para que os requisitos sejam escritos em LN [15] e permaneçam nessa linguagem: é uma

forma de comunicação primária compartilhada por todos os atores do processo de desenvolvimento de software; os requisitos são especificados considerando diferentes níveis de abstração; a relação custo benefício, a necessidade de reação às mudanças de mercado, a rapidez necessária ao desenvolvimento do produto de software, são favoráveis à LN; oferece melhor apoio para a gestão de requisitos errados, incompletos ou parcialmente definidos; pode capturar propriedades externas dos requisitos, como as reais intenções dos usuários.

Além disso, representações em linguagem natural, segundo [16], tendem a ser mais fáceis de entender para todos os stakeholders sendo, portanto, mais acessíveis. Para [17], os requisitos são especificados, na maior parte da indústria, utilizando uma linguagem natural não estruturada. Segundo esse autor, é mais fácil para todos os interessados, mesmo aqueles que carecem de formação técnica, entender a linguagem natural do que outras notações não textuais.

Já [18], em um estudo experimental realizado na Argentina sobre práticas de Engenharia de Requisitos no desenvolvimento de aplicações web em 25 empresas de software, analisaram, entre outros aspectos, o tipo de representação dos requisitos. Os resultados indicaram que 52% da amostra utilizam linguagem natural e 64% utilizam casos de uso, sendo que algumas empresas utilizam mais de uma representação.

Por outro lado, os requisitos escritos em LN podem ser imprecisos, incompletos e ambíguos. A informalidade pode ajudar na discussão entre as partes interessadas no início do projeto, mas pode levar à confusão, inconsistência, falta de automação, ambiguidade e aos erros [19] [20]. Outro desafio na especificação dos requisitos de software, utilizando a LN, é o grande volume de documentos gerados, notadamente em grandes projetos. A medida que cresce a documentação da especificação, aumentam as dificuldades com a interpretação, a consistência e a manutenção [5].

Para reduzir a ambiguidade, uma das soluções foi o desenvolvimento de um Léxico Ampliado da Linguagem [21] tornando mais preciso o vocabulário utilizado na especificação de requisitos. Segundo [19], técnicas de Processamento de Linguagem Natural de têm sido propostas para melhorar a especificação dos requisitos utilizando o modo semi-automático ou totalmente automático, mas até agora não têm sido amplamente implementadas. Alguns pesquisadores afirmam que o processamento de linguagem não é maduro o suficiente para ser aplicado em engenharia de requisitos, mas várias propostas têm sido apresentadas com resultados promissores.

Autores com [22] investigaram o estado atual da automação do processo de elicitação de requisitos, em linguagem natural, considerando as pesquisas existentes que identificam ferramentas automatizadas e semi-automatizadas a partir de documentos de elicitação de requisitos. Realizaram uma revisão sistemática considerando 36 artigos e classificaram esses trabalhos em grupos de ferramentas de total e parcial automatização, com relação a vários domínios, entre eles a elicitação de requisitos. Segundo esses autores, a Linguagem

Natural com sua ambiguidade e inconsistência inerentes, força os engenheiros a verificar os requisitos manualmente e, por vezes, os resultados do automatismo são o retrabalho. Isto inclui corrigir erros, estender ou limitar as passagens de texto identificado ou simplesmente adicionando requisitos que foram desconsiderados. Consequentemente, o automatismo geralmente é complementado com funcionalidades manuais de suporte a elicitação.

Reforçando essa afirmação, em [23] se propõem a adaptar os avanços recentes em processamento de linguagem natural e tecnologias semânticas para gerar modelos UML, teste de casos, e mesmo código fonte. Ao final sugerem a necessidade de evolução de classes de ferramentas, estudo ontológico e outros aspectos para que isso seja possível, apesar de alguns resultados promissores de seu trabalho. Citam também a necessidade de interferência humana para resolver problemas de ambiguidades.

Considerando o contexto de elicitação de requisitos não funcionais, autores como [24], citam que a detecção precoce de requisitos não-funcionais é crucial na avaliação da arquitetura. Esses autores propõem uma abordagem de categorização textual semi-supervisionado para a identificação automática e classificação dos requisitos não-funcionais. Utilizam um método para detecção e classificação e técnicas de aprendizado semi-supervisionado. Essa abordagem semi-supervisionada resultou em taxas de acerto acima de 70%, usando coleções padrões de documentos.

Ou seja, utilizando ou não modelos automatizados, semi-automatizados ou manuais, as organizações continuam a utilizar a linguagem natural em parte das atividades na engenharia de software, especificamente na especificação de requisitos. Identificar e compreender a percepção dos usuários com relação a essa utilização pode vir a contribuir para esse processo no futuro. A seguir detalha-se a metodologia da pesquisa.

III. METODOLOGIA DA PESQUISA

O método adotado neste estudo foi o descritivo, de abordagem quantitativa por meio de equações estruturais. Equações estruturais é um termo que não designa uma técnica estatística específica, mas a uma série de técnicas e procedimentos utilizados para validar modelos teóricos que definem relações causais, hipotéticas entre variáveis [25], explicitando através de parâmetros o efeito da interação entre elas.

O modelo teórico proposto baseia-se na UTAUT. Esse modelo foi elaborado para explicar a aceitação de uma tecnologia por um indivíduo ou grupo [26]. Por ser comprovado e bem aceito em pesquisas científicas [27], ele foi escolhido para servir de base para esse trabalho.

Para isso elaboram-se uma série de hipóteses a partir do modelo UTAUT e usa-se as respostas de um questionário para confirmá-las ou rejeitá-las através de fórmulas matemáticas.

Por meio da base teórica e segundo o modelo, existem quatro fatores independentes que poderão influenciar a aceitação de uma tecnologia: Expectativa de Desempenho (PE), Expectativa de Esforço (EE), Influência Social e

Condições Facilitadoras (FC), e dois fatores dependentes destes: Intenção de Uso (BI) e Comportamento de Uso (IC), conforme se vê na Fig. 1.

Seguindo esses fatores e para responder aos questionamentos da pesquisa, foram elaboradas hipóteses adaptadas para a aceitação e uso da Linguagem Natural.

- H1: A expectativa de um rendimento elevado leva os analistas a usar a Linguagem Natural na Especificação de Requisitos.
- H2: A expectativa de um baixo esforço leva os analistas a usar a Linguagem Natural na Especificação de Requisitos.
- H3: O uso de Linguagem Natural na Especificação de Requisitos leva a uma maior adesão às necessidades do cliente.
- H4: A organização onde o analista trabalha estimula o uso da Linguagem Natural na Especificação de Requisitos.

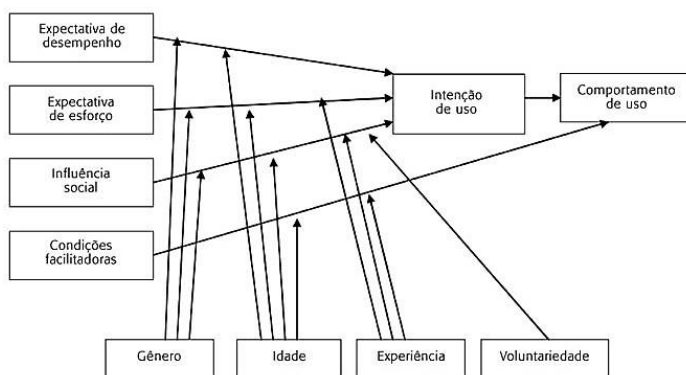


Fig. 1. O modelo UTAUT original [28]

O modelo UTAUT foi adaptado às hipóteses, obtendo-se o resultado apresentado na Fig. 2. A Influência Social foi descartada devido a que o uso da LN não é obrigatório [27].

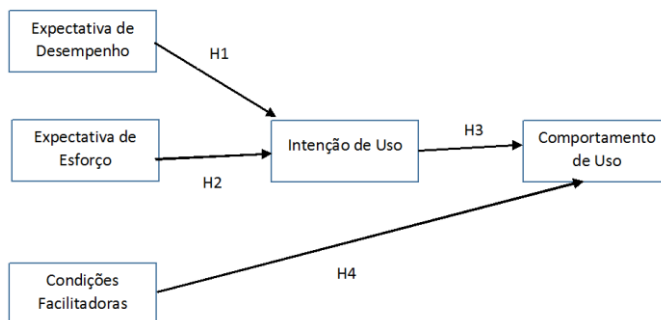


Fig. 2. Modelo UTAUT adaptado

O modelo é formado por variáveis que se apresentam em diversas ocasiões com aspectos subjetivos, necessitando de indicadores objetivos para realizar uma aproximação ao seu conceito. Deste modo, a técnica apropria-se de desenhos experimentais ou quase experimentais por meio de questionários ou bases de dados [27]. Assim, foi utilizado um questionário como input dos dados. O instrumento foi

adaptado do próprio modelo de UTAUT, com 17 questões relacionadas à expectativa de desempenho, expectativa de esforço, condições facilitadoras, intenção de uso e comportamento no uso da LN, objetivando identificar a aceitação e uso da LN nas Especificações de Requisitos pelos utilizadores dessa tecnologia (Quadro 1).

O questionário foi disponibilizado via *Google Docs* e enviado a profissionais que atuam com análise de requisitos em empresas do ramo de tecnologia (indústria de software) na Argentina e no Brasil [10]. Os participantes acessaram uma URL indicada e responderam a questões relacionadas indicando o grau de concordância com as afirmações expostas (uso da escala de Likert de 7 pontos): 1 para discordo inteiramente, 2 para discordo em grande parte, 3 para discordo parcialmente, 4 para neutro, 5 para concordo parcialmente, 6 para concordo em grande parte e 7 para concordo inteiramente.

QUADRO 1
FATORES E INDICADORES DO MODELO UTAUT ADAPTADO À TECNOLOGIA LN NAS ESPECIFICAÇÕES DE REQUISITOS

Expectativa de Desempenho
PE1 - A linguagem natural é útil na especificação de requisitos.
PE2 - Utilizar a linguagem natural capacita-me a executar a especificação de requisitos mais rapidamente.
PE3 - Utilizar a linguagem natural aumenta minha produtividade na especificação de requisitos.
PE4 - Se eu uso a linguagem natural eu aumento minhas chances de ganhar uma promoção.
Expectativa de Esforço
EE1 - Minha interação com a linguagem natural é clara e compreensiva.
EE2 - É fácil tornar-se experiente em usar linguagem natural para especificar requisitos.
EE3 - É fácil usar a linguagem natural para especificar requisitos.
EE4V - É fácil aprender a usar a linguagem natural.
Condições Facilitadoras
FC1 - Eu tenho os recursos necessários para usar a linguagem natural na especificação de requisitos.
FC2 - Eu tenho o conhecimento necessário para usar a linguagem natural na especificação de requisitos.
FC3 - A linguagem natural não é compatível com outros métodos que eu uso para especificar requisitos.
FC4 - Uma pessoa específica (ou grupo) está disponível para prover-me assistência nas dificuldades com a linguagem natural.
Intenção de Uso da Linguagem Natural
B11 - Eu pretendo continuar usando a linguagem natural para especificar requisitos no futuro.
B12 - Eu sempre tento usar a linguagem natural na especificação de requisitos.
B13 - Eu planejo continuar usando a linguagem natural frequentemente na especificação de requisitos.
Comportamento no uso da Linguagem Natural
IC1 - Eu já fiz uso da linguagem natural na especificação de requisitos.
IC2 - Eu uso a linguagem natural na especificação de requisitos com muita frequência.

IV. RESULTADOS E ANÁLISE

Os dados obtidos pelo questionário disponibilizado via *Google Docs* na Argentina, referem-se ao período de outubro e dezembro de 2016 e no Brasil [10], ao período de maio a outubro de 2015. Foram respondidos 45 questionários na Argentina e 50 no Brasil. A amostra na Argentina foi constituída em 69% por homens e 31% por mulheres, todos com idade média de 34 anos e 10 meses, e com uma

experiência em requisitos média de 7 anos e 6 meses. No Brasil, identificou-se que 66% dos analistas respondentes eram homens e 34% mulheres, todos com idade média de 37 anos e 3 meses. A amostra apresentava uma média de experiência com requisitos de 7 anos e 3 meses.

A. Argentina

O modelo proposto foi submetido a testes com a finalidade de dotá-lo de confiabilidade e validade no contexto Argentino. Após a coleta de dados, estes foram consolidados em uma planilha Excel. Foram usados os de confiabilidade simples, confiabilidade composta, variância explicada e validade discriminante, utilizando-se o software SmartPLS (*Smart Partial Least Square*) 2.0. O software é uma ferramenta para modelagem de equações estruturais [28] e para a análise de experimentos que exijam múltiplas variáveis ou mesmo relacionar influência direta e indireta de diversos fatores. Os modelos de equações estruturais utilizam modelos estatísticos de múltiplas variáveis para avaliar o grau de predição de variáveis independentes sobre as variáveis dependentes. Porém inicialmente é necessário saber se os itens referentes a cada variável estão correlacionados entre si (Tabela I).

TABELA I
CONFIABILIDADE DE ITEM ANTES E DEPOIS DA DEPURAÇÃO - ARGENTINA

Variáveis	Itens	Confiabilidade de Item	Confiabilidade de Item depurado
Expectativa de Desempenho	PE1	0,56	0,56
	PE2	0,48	0,48
	PE3	0,90	0,91
	PE4	0,64	0,72
Expectativa de Esforço	EE1	0,84	0,89
	EE2	0,86	0,81
	EE3	0,65	0,65
	EE4	0,45	0,45
Condições Facilitadoras	FC1	0,99	1,00
	FC2	0,99	1,00
	FC3	-0,28	-0,28
	FC4	0,22	0,22
Intenção de Uso da Linguagem Natural	B11	0,74	0,73
	B12	0,88	0,88
	B13	0,93	0,93
Utilidade na Especificação de Requisito	IC1	0,96	0,96
	IC2	0,97	0,97

Para assegurar a confiabilidade de item, os valores de correlação devem ser iguais ou superiores a 0,707 [28]. Pode-se observar que os indicadores PE1, PE2, EE3, EE4, FC3 e FC4 não foram confiáveis, assim foram depurados do modelo. A cada retirada de item os valores são recalculados.

Uma vez realizada a confiabilidade dos itens, os próximos testes a serem realizados são confiabilidade composta e os

testes de validade como variância média extraída AVE (*Average Variance Extracted*) e Validade Discriminante.

Para assegurar um valor aceitável para confiabilidade composta, é necessário que os valores sejam iguais ou superiores a 0,7 [28]. A aprovação deste teste garante que o grupo de itens de cada variável seja suficiente para explicá-las. O primeiro teste de validade é o AVE, responsável por explicar o grau de conversão dos itens para sua respectiva variável e não para outra, pois um indicador que não possua ao menos 50% de diferença em relação aos indicadores de outras variáveis, pode ocasionar em um problema de validade. Assim, para valores de AVE satisfatórios, espera-se valores maiores ou iguais a 0,5 ou 50% [28]. Na Tabela II encontram-se os valores do modelo Argentino.

TABELA II
CONFIABILIDADE COMPOSTA E AVE - ARGENTINA

Variáveis	Confiabilidade composta	AVE
Condições Facilitadoras	1,00	1,00
Expectativa de Desempenho	0,80	0,68
Expectativa de Esforço	0,85	0,74
Intenção de Uso da Linguagem Natural	0,89	0,73
Utilidade na Especificação de Requisitos	0,96	0,93

O último teste de validade é a Validade Discriminante que assegura que cada variável tenha independência em relação às demais. Para assegurar sua autonomia no modelo se deve garantir que a raiz quadrada do AVE seja superior as forças das correlações (Tabela III).

TABELA III
VALIDADE DISCRIMINANTE – ARGENTINA

	Condições Facilitadoras	Expectativa de Desempenho	Expectativa de Esforço	Uso da Linguagem Natural
Condições Facilitadoras	1			
Expectativa de Desempenho	0,324569	0,8216337		
Expectativa de Esforço	0,468128	0,472066	0,858419	
Intenção de Uso da Linguagem Natural	0,498549	0,580291	0,485742	0,851697
Utilidade na Especificação de Requisitos	0,660265	0,299634	0,438285	0,589262

As raízes quadradas aparecem (Tabela III) em negrito e seu valor deve ser superior à força das correlações das demais variáveis que estão abaixo de cada raiz quadrada em negrito. Pode-se observar que as raízes quadradas de AVE foram superiores às cargas fatoriais dos itens, assim pode-se dizer que o modelo é válido e confiável, podendo ser usado em

pesquisas futuras que adotem o mesmo instrumento.

Os resultados permitem a validação do modelo proposto, permitindo avaliar as hipóteses propostas (Fig. 3).

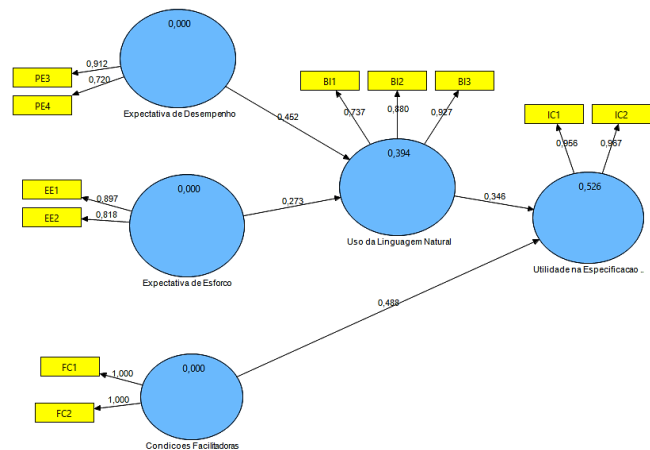


Fig. 3. Modelo de pesquisa depurado

Uma vez atestada a confiabilidade e validade do modelo, segue-se para a mensuração e resultados na análise. A valoração do modelo é realizada através da Análise do R2 e do Beta. O R2 explica em que grau a variável dependente é predita pela independente. E o Beta explica a força de influência de cada variável independente sobre a dependente.

Pode-se observar na Tabela IV que a Intenção de Uso da Linguagem Natural é explicada pela Expectativa de Esforço e pela Expectativa de Desempenho em um 39,4% (R2 = 0,394), enquanto a Utilidade na Especificação de requisitos foi predito pelas Condições Facilitadoras e da Intenção de Uso da Linguagem Natural, que é o problema central deste estudo, em 52,6% (R2 = 0,526).

TABELA IV
VALORAÇÃO DO MODELO – ARGENTINA

Hipóteses	Beta	T Statistics	R2
H1-Expectativa de Desempenho → Intenção de Uso de Linguagem Natural	0,452	4,85	0,394
H2-Expectativa de Esforço → Intenção de Uso de Linguagem Natural	0,273	2,88	
H3-Intenção de Uso da Linguagem Natural → Utilidade na Especificação de Requisitos	0,346	3,66	0,526
H4-Condições Facilitadoras → Utilidade na Especificação de Requisitos	0,488	4,64	

Para que seja aceitável a predição deve ser de ao menos 10%, sendo reveladora acima de 30% [27]. Neste caso pode-se dizer que o estudo tem uma predição reveladora.

Analisando os Betas, pode-se perceber (Tabela IV) a expectativa de desempenho (0,452) possui uma maior influência na Intenção de Uso da linguagem Natural do que a Expectativa de Esforço (0,273), e a Utilidade na Especificação de Requisitos é influenciada com mais força pelas Condições

facilitadoras (0,488) do que pela Utilidade na Especificação de Requisitos (0,346). Os Betas revelam o primeiro teste de hipótese e para que as hipóteses sejam válidas devem apresentar um valor igual ou superior a 0,2 [27].

Um segundo teste de hipóteses é o *T-Student*. Pode-se perceber que o teste *T-Student* (Tabela IV), são superiores a 1,96 [28], confirmando todas as hipóteses do modelo argentino como verdadeiras.

Assim pode-se perceber que a Utilidade na especificação de requisitos apenas é mais efetiva se existem condições facilitadoras para tal, que se pode traduzir desde apoio da alta direção da empresa, ou mesmo colaboração da equipe envolvida. Outro fator a ser levado em consideração é a expectativa do desempenho que impacta fortemente a intenção de uso da Linguagem Natural, ou seja, será usada se percebe uma melhoria potencial do desempenho. Sem embargo, todos os fatores se mostraram importantes, porém o maior grau de importância das condições facilitadoras e da expectativa de desempenho norteiam quais fatores devem ser trabalhados como prioridade. Assim todas as hipóteses foram aceitas como verdadeiras.

B. Brasil

No ano anterior, um estudo foi realizado por parte da equipe, tendo como alvo o Brasil. A seguir apresentam-se sucintamente os resultados do Brasil [10]. A Tabela V apresenta o modelo mensurado elaborado para este país. Os critérios de confiabilidade e validade foram realizados e obtiveram os resultados desejados para o modelo brasileiro.

Pode-se observar na Tabela V que a LN é explicada pela Expectativa de Esforço e pela Expectativa de Desempenho em 14,7%, enquanto o Comportamento de Uso da LN na Especificação de Requisitos, problema central do estudo brasileiro, foi predita em 63,4% pelas variáveis Intenção de Uso da Linguagem Natural e Condições Facilitadoras. Para que seja aceitável a predição deve ser de ao menos 10%, sendo reveladora acima de 30% [28]. Neste caso pode-se se dizer que o estudo tem uma predição reveladora.

TABELA V
VALORAÇÃO DO MODELO – BRASIL

Hipóteses	Beta	T Statistics	R2
H1-Expectativa de Desempenho → Intenção de Uso de Linguagem Natural	0,296	0,60	0,147
H2-Expectativa de Esforço → Intenção de Uso de Linguagem Natural	0,138	2,38	
H3-Intenção de Uso da Linguagem Natural →Utilidade na Especificação de Requisitos	0,776	0,92	0,634
H4-Condições Facilitadoras → Utilidade na Especificação de Requisitos	0,062	9,08	

Embora os valores de predição sejam satisfatórios, apenas duas hipóteses foram validadas parcialmente (H1 e H3), apresentando betas superiores a 0,2, porém falhando no teste *T-Student*.

C. Comparação entre a Argentina e o Brasil

A seguir apresentam-se os graus de influência de cada variável independente em sua dependente. Esta análise é desvelada através dos valores de Beta. A literatura explica que para que as relações que se apresentam como hipóteses devem ter valor igual ou superior a 0,2 para serem significativas e apenas as relações significantes asseguram hipóteses verdadeiras [28].

Assim, para os dados argentinos, o grau de influência de cada variável nas predições (Beta) obtidas foram respectivamente de 0,468, de 0,254, de 0,346 e 0,488, podem dar confiabilidade à comprovação das hipóteses H1, H2, H3 e H4. Com relação aos dados brasileiros, o grau de influência de cada variável obtida foi respectivamente de 0,307, de 0,121, de 0,776 e de 0,062. Assim, para os dados brasileiros somente se pode dar confiabilidade à comprovação das hipóteses H1 e H3.

Os resultados dos dois países permitem constatar o reconhecimento por parte dos analistas argentinos e brasileiros que o uso da LN leva a maior adesão às necessidades dos clientes (H3). Usando LN é provável que a comunicação entre as partes se torne mais transparente, aumentando a qualidade dos requisitos. Isso ratifica os achados citados em outras pesquisas [15].

Os resultados permitem também identificar que há uma forte concordância dos analistas argentinos e brasileiros de que a Expectativa de Desempenho elevado conduz à Intenção de Uso de Linguagem Natural na especificação de requisitos (H1). Ou seja, com a melhor troca de informações entre as partes e, assim, evitando ruídos na comunicação é possível levantar os mesmos requisitos em menos tempo [19].

Com índices menores de concordância, mais ainda significativos, os dados argentinos permitem identificar que: a Expectativa de um baixo esforço leva os analistas a usar a Linguagem Natural na Especificação de Requisitos (H2); e que as organizações onde os analistas trabalham estimulam o uso da LN na especificação de requisitos (H4). Esses índices não são significativos nos dados brasileiros. Isso possibilita inferir que, talvez a especificação formal de modelos para os requisitos, nas organizações brasileiras, pode estar sendo preferida para retratar os mesmos conceitos em um nível mais alto de abstração, facilitando a visualização do problema e da sua solução, a rastreabilidade entre os diversos modelos e a aproximação com a fase de desenvolvimento do software [15].

Pode-se também mensurar o percentual de influência destas variáveis dos resultados argentinos e brasileiros multiplicando a o Beta da Variável Latente por suas Correlações (Tabela VI).

Nesta análise, com relação aos dados argentinos, verifica-se que as Condições Facilitadoras se destacaram, colaborando em 32,2% na predição da Intenção de Uso da Linguagem Natural, assim como a Expectativa de Desempenho influencia em 27,2% na Intenção de Uso da Linguagem Natural. No caso brasileiro, a Expectativa de Desempenho também influencia em 11,4% na intenção de uso da linguagem natural. Nos dados brasileiros a Utilidade na especificação é a variável que mais influencia (61,6%) a intenção de uso da linguagem natural.

TABELA VI
PERCENTUAL DE INFLUÊNCIA DAS VARIÁVEIS ARGENTINA VS BRASIL

Variáveis	Argentina				Brasil			
	BI	IC	Correlação das variáveis latentes	% de Explicação	BI	IC	Correlação das variáveis latentes	% de Explicação
FC		0,488	0,6603	32,22%		0,062	0,283	1,8%
EE	0,273		0,4857	13,26%	0,121		0,280	3,4%
PE	0,452		0,5803	26,23%	0,307		0,370	12,4%
BI		0,346	0,5889	20,38%		0,776	0,794	61,6%

V. CONCLUSÃO

O objetivo deste trabalho foi realizar um estudo comparativo sobre os fatores de influência para a intenção de uso da LN na especificação de requisitos considerando a percepção de técnicos de dois países: Argentina e o Brasil. A metodologia adotada foi um estudo descritivo, de abordagem quantitativa por meio de equações estruturais.

O modelo teórico proposto baseia-se na UTAUT. Foi elaborado um questionário, validado com $\alpha=0,817$, com 17 perguntas para 45 analistas de requisitos da Argentina e 50 analistas de requisitos do Brasil que trabalham na indústria de software.

O instrumento de pesquisa foi validado e pode continuar a ser usado.

Identificou-se que os motivos que levam os analistas dos dois países a utilizar a LN na especificação é a expectativa de um rendimento elevado (H1) e a percepção que o uso da LN leva a uma maior adesão às necessidades do cliente (H3). Dessa forma as Hipóteses 1 e 3 foram comprovadas para ambos os países. São elas:

- H1: A expectativa de um rendimento elevado leva os analistas a usar a Linguagem Natural na Especificação de Requisitos.
- H3: O uso de Linguagem Natural na Especificação de Requisitos leva a uma maior adesão às necessidades do cliente.
- As outras expectativas foram confirmadas com fatores de influência para um dos países participantes da pesquisa (Argentina). Ou seja, para a Argentina as Hipóteses 2 e 4 também foram confirmadas.

Destaca-se que essas hipóteses não foram confirmadas para o Brasil. O que permite inferir que, no Brasil, apesar dos analistas de requisitos reconhecerem a importância do uso da LN. Por outro lado, os mesmos profissionais não identificam condições facilitadoras para sua utilização mais efetiva nas organizações.

Essa constatação permite, portanto, verificar que, a despeito do embasamento técnico e conceitual fornecido pelas pesquisas nos últimos anos, é preciso a realização de investigações científicas voltadas o apoio organizacional sobre a utilização da LN na especificação dos requisitos. Estudos futuros poderiam apontar quais são as razões para esse

descompasso entre a visão técnica e a visão gerencial.

Sugere-se ainda, a expansão da pesquisa para outros países da América Latina de modo a se obter uma melhor compreensão desse mercado, conforme forem sendo comprovados ou ratificados os achados atuais.

REFERÊNCIAS

- [1] E. J.R. de Castro, A.T.S. Calazans, R. A. Paldes, F. de A. Guimaraes, *Engenharia de Requisitos Um enfoque prático na construção de software orientado ao negócio*, Florianópolis: Bookess, 2014.
- [2] J. A. Dorigan, and R. M. de Barros, "A model of requirements engineering process for standardization and quality increase", in *Proc. of ADIS International Conference Applied Computing*, 2012, pp.343-347.
- [3] I. Sommerville, *Engenharia de software*, 9.ed., Addison-Wesley, 2011.
- [4] R.. G. Bustos, "Procesamiento de Lenguaje Natural en Ingeniería de Requisitos: Contribuciones Potenciales y Desafios de Investigación", in *Proc. of CIBSE 2015: Conferencia Iberoamericana de Software Engineering*, Lima, 2015, pp.1-3.
- [5] P. S. Lopes, "Uma taxonomia da pesquisa na área de Engenharia de Requisitos", Tese de Mestrado, Universidade de São Paulo, 2002.
- [6] E. Kroth, and G. Dessbesel, "Emprego de técnicas de representação do conhecimento como forma de apoio à engenharia de requisitos", in *Proc. of XVIII Congreso Argentino de Ciencias de la Computación*, 2012.
- [7] P. Ramirez-Correa, "Uso de internet móvil en Chile: explorando los antecedentes de su aceptación a nivel individual", *Ingeniare. Revista chilena de ingeniería*, vol. 22, no. 4, pp. 560-566, 2014.
- [8] J. Farias, P. Lins, P. Albuquerque, "A propensão de usuários à adoção de tecnologias: um estudo com usuários e não usuários do Progama Nota Legal no Distrito Federal", in *Proc. of XI Simpósio Brasileiro de Sistemas de Informação*, Goiânia, 2015.
- [9] M. Visentini, F. Chagas, D. Bobsin, "Análise Bibliométrica das Pesquisas sobre redes sociais virtuais publicadas em âmbito nacional", in *Proc. of SEPE-Seminário de Ensino, Pesquisa e Extensão da UFFS*, 2015.
- [10] R.A. Paldês, A.T.S. Calazans, A.M. Mariano, E.J.R.de Castro, "A utilização da linguagem natural na especificação de requisitos: um estudo por meio das equações estruturais", in *Proc. of XIX Ibero-American Conference on Software Engineering*, 2016, pp. 365-378.
- [11] SWEBOOK. *The Guide to the Software Engineering Body of Knowledge*, Los Alamitos, California: IEEE Computer Society, 2014.
- [12] B. Manrique Losada, and C. Zapata Jaramillo, "Transformación de lenguaje natural a controlado en la educación de requisitos: una síntesis conceptual basada en esquemas preconceptuales", *Revista Facultad de Ingeniería de Antioquia*, no. 70, pp.132-145, Jan/Mar 2014.
- [13] V. M. A. Lima, "Terminologia, comunicação e representação documentária", Tese de Mestrado, Universidade de São Paulo, 1998.
- [14] IEEE 29148-2011, IEEE Systems and software engineering - Life cycle processes - Requirements engineering. IEEE, Nova York, 2011.
- [15] M. Sayão, "Verificação e validação de requisitos: Processamento da Linguagem natural e agentes", Tese de Doutorado, Pontificia Universidade Católica do Rio de Janeiro, 2007.
- [16] A. Sutcliffe, and P. Sawyer, "Requirements Elicitation: Towards the Unknown Unknowns", in *Proc. of Research Track IEEE, RE 2013*, Rio de Janeiro, 2013.
- [17] J. M. C. Geal, J. Nicolas, J. L. F. Aleman, A. Toval, S. Ouhbi, and A. Idri, "Co-located and distributed natural-language requirements specification: traditional versus reuse-based techniques", *Journal of software: evolution and process*, vol. 28, no. 3, pp.205-227, 2016.
- [18] A. Oliveros, F. J. Danyans, and M. L. Mastropietro, "Prácticas de Ingeniería de Requerimientos en el desarrollo de aplicaciones Web", in *Proc. of CIBSE 2014: track Workshop em Engenharia de Requisitos*, 2014.
- [19] S. Ghosh, D. Elenius, W. Li, P. Lincoln, N. Shankar, and W. Steiner, "ARSENAL: Automatic Requirements Specification Extraction from Natural Language", Cornell University Library, Jul., p.p.1-32, 2014.
- [20] H. Yang, , A. Roeck, V. Gervasi, A. Willis., and B. Nuseibeh, "Analysing Anaphoric Ambiguity in Natural Language Requirements", *Requirements Eng*, vol. 16, no. 3, pp. 163-189, May 2011.
- [21] P. Engiel, J. Pivatelli, P. Moura, R. Portugal, and J. C. S. P. Leite, "Um processo colaborativo para a construção de léxicos: o caso da divulgação de transparência", in *Proc. of CIBSE 2015: track XVIII Workshop on Requirements Engineering*, Lima, Perú, 2015.

- [22] H. Meth, and A. Maedche, “The State of the Art in Automated Requirements Elicitation”, *Information and Software Technology*, October 2013.
- [23] W. F. Tichy, and S. J. Koerner, “Text to software: developing tools to close the gaps in software engineering”, in *Proc. of the FSE/SDP Workshop on Future of Software Engineering Research (FoSER'10)*, 2010, pp. 379–384.
- [24] A. Casamayor, D. Godoy, and M. Campo, “Identification of non-functional requirements in textual specifications: a semi-supervised learning approach”, *Information and Software Technology*, vol. 52, no. 4, pp. 436–445, 2010.
- [25] J. Marôco, “Análise de Equações Estruturais: Fundamentos teóricos, software & aplicações”, *ReportNumber*, Pêro Pinheiro, 2010.
- [26] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, “User acceptance of information technology: Toward a unified view”, *MIS Quarterly*, vol. 27, no. 3, pp. 425-478, 2003.
- [27] P. E. Ramirez, A. M. Mariano, and E. A. Salazar, “Propuesta Metodológica para aplicar modelos de ecuaciones estructurales con PLS: El caso del uso de las bases de datos científicas en estudiantes universitarios”, *Revista ADMpg Gestão Estratégica*, vol. 7, no. 2, artículo 15, 2014.
- [28] W.W. Chin, “The partial least squares approach for structural equation modelling, en Methodology for Business and Management”, *Modern Methods for Business Research*, p.p. 295-336, Hillsdale: Lawrence Erlbaum Associates, 1998.