

RAE – CEA – 17P10

**RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO: "O ESTATUTO
TEÓRICO DO SUJEITO NULO DO PORTUGUÊS BRASILEIRO: RESULTADOS
EXPERIMENTAIS"**

Anatoli Yambartsev

Yangyang Chen

CENTRO DE ESTATÍSTICA APLICADA – CEA – USP
RELATÓRIO DE ANÁLISE ESTATÍSTICA

TÍTULO: Relatório de análise estatística sobre o projeto: "O estatuto teórico do sujeito nulo do português brasileiro: resultados experimentais"

PESQUISADORAS: Claudia Souza Coelho
Letícia Evelyn Leite Santos

ORIENTADOR: Jairo Nunes

INSTITUIÇÃO: Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paul – FFLCH-USP

FINALIDADE: Iniciação científica

RESPONSÁVEIS PELA ANÁLISE: Prof. Dr. Anatoli Yambartsev
Yangyang Chen

REFERÊNCIA DESTE TRABALHO:

YAMBARTSEV, A., CHEN, Y. (2017) **Relatório de análise estatística sobre o projeto "O estatuto teórico do sujeito nulo do português brasileiro: resultados experimentais"** São Paulo, IME – USP, RAE – CEA – 17P10

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS:

BUSSAB, W.O. E MORRETTIN, P.A. **Estatística Básica**. 8ª ed., 5ª tiragem São Paulo: Saraiva, 2013

FERRARI, S. L. P. e CRIBARI-NETO, F. **Beta Regression for Modelling Rates and Proportions**, Journal of Applied Statistics, 2004.

SMITHSON, M. e MERKLE, E. C. **Generalized Linear Models for Categorical and Continuous Limited Dependent Variables**, CRC Press, 2014

PROGRAMAS COMPUTACIONAIS UTILIZADOS

Microsoft Excel 2016 para Mac;

Microsoft Word 2016 para Mac;

R (versão 3.3.1);

RStudio (versão 1.0.143)

TÉCNICAS ESTATÍSTICA UTILIZADAS

Análise Descritiva Unidimensional (03:010);

Análise Descritiva Multidimensional (03:020);

Testes de Hipóteses Paramétricas (05:010);

Testes de Hipóteses Não Paramétricas (05:070);

Análise de Regressão Clássica (07:020)

ÁREA DE APLICAÇÃO

Linguística (14:110)

Sumário

Resumo	5
1. Introdução.....	6
2. Descrição do projeto.....	7
2.1. Descrição do estudo	7
2.1. Descrição das variáveis.....	8
3. Exemplos	9
4. Análise descritiva	10
4.1. Sentenças dos tipos <i>familiarização, distratora e convencer</i>	10
4.2. Sentenças dos tipos <i>adjunto, complemento e convencer</i>	11
5. Análise inferencial.....	14
5.1. Testes de independências.....	14
5.2. Testes de correlação.....	14
5.3. Testes de tendência	15
5.4. Modelo de regressão beta	15
Conclusão	18
Anexo 1 - As sentenças do teste.....	19
Apêndice A – Tabelas da Análise Descritiva.....	22
Apêndice B – Gráficos da Análise Descritiva	35
Apêndice C – Tabelas da Análise Inferencial.....	50
Apêndice D – Diagnóstico do Modelo.....	53

Resumo

O português brasileiro é uma língua que possui sujeito nulo e, existem duas análises para identificar como o sujeito nulo referencial do português brasileiro se comporta. O problema é que não há consenso sobre o estatuto teórico dele. Neste projeto, o objetivo é analisar o comportamento de sentenças que possuem o verbo "convencer" a fim de obter evidências para o estudo do sujeito nulo.

Da análise descritiva e da análise inferencial neste estudo concluiu-se que as respostas dos falantes nativos indicam que maioria deles consideraram que as sentenças do tipo *convencer* têm função de complemento. A estrutura das respostas das sentenças do tipo *convencer* parece mais com a do tipo *adjunto* nas sentenças do tipo sintagma interrogativo simples e parece mais com a do tipo *complemento* nas sentenças do tipo de sintagma interrogativo complexo. O fator ordem não interfere nas respostas das sentenças do tipo *convencer*. O modelo de regressão beta ajustado, em que a variável resposta é a proporção de acertos das sentenças, indica que só o tipo de sentença é significativa.

1. Introdução

O sujeito, gramaticalmente, é um dos termos essenciais da oração, responsável por realizar ou sofrer uma ação ou estado e encontra-se, geralmente, à esquerda do verbo. Um dos tipos de sujeito é o que não está foneticamente expresso na oração, chamado de sujeito nulo. As orações com sujeito nulo serão sempre licenciadas por alguma outra informação para que se possa identificar a que pessoa do discurso a sentença se refere. Há duas diferentes formas para identificar os sujeitos nulos: pela desinência¹ do verbo ou pelo contexto em que a sentença está inserida.

O português brasileiro é uma língua tipicamente de sujeito nulo, pois apresenta propriedades sintáticas de línguas pro-drop². Além do português brasileiro, há muitas outras línguas que possuem sujeito nulo, como espanhol, italiano e português europeu, porém, suas propriedades sintáticas são fundamentalmente diferentes.

Existem duas análises para identificar como o sujeito nulo referencial do português brasileiro se comporta de forma diferente do de outras línguas pro-drop: a primeira propõe que esse sujeito é pronome nulo defectivo e a segunda propõe que é um vestígio de movimento sintático. O problema é que as condições de legitimação desse sujeito em cada uma das abordagens são muito semelhantes, ou seja, não há consenso sobre o estatuto teórico dele. Logo, a tarefa de distingui-las empiricamente é bastante difícil, e os julgamentos de gramaticalidade³ por falantes nativos de português brasileiro não são tão claros.

Nesse contexto, um dos principais pontos levantados contra a análise de movimento adotada por este projeto é o comportamento do sujeito nulo do português brasileiro em sentenças com o verbo "convencer". O experimento busca esclarecer, com seus resultados, divergências de julgamento encontradas na literatura e apontar qual análise do sujeito nulo do português brasileiro é a mais adequada.

¹ Desinências são os elementos terminais indicativos das flexões das palavras.

² Redução de pronoun-dropping, do inglês para 'supressão de pronome'.

³ O falante nativo de uma língua tem intuição sobre o que é uma sentença "normal" ou não na língua dele.

2. Descrição do projeto

2.1. Descrição do estudo

Os dados foram coletados e categorizados com o auxílio do software de TP⁴, que é um software gratuito com versão em português, desenvolvido para realização de experimentos de percepção da fala e permite estímulos de áudio e/ou imagem. O grupo de participantes do estudo foi composto por dez pessoas do gênero masculino e dez pessoas do gênero feminino (uma participante foi excluída, pois respondeu SIM para todas as sentenças), todos falantes nativos do português brasileiro nascidos, criados no Estado de São Paulo e que têm entre vinte e trinta anos.

Os participantes foram informados de que estava sendo avaliado o desempenho de um software de computador na formulação de perguntas complexas em português e, para isso, se precisava que falantes da língua julgassem quais perguntas formuladas pelo software eram bem formadas e quais não eram. Os participantes ouviam gravações das sentenças relevantes lidas por um falante de português e pressionavam as teclas SIM ou NÃO para dizer se as sentenças eram ou não bem formadas. Os participantes tinham também a possibilidade de pressionar a tecla REPETIR se desejassem ouvir a sentença uma vez mais antes de pressionar SIM ou NÃO.

O processo de coleta de dados foi precedido por uma familiarização, em que os participantes foram expostos a quatro pares de sentenças gramaticais e agramaticais, para que se acostumassem à tarefa que deveriam realizar. Depois, foram expostos a oito sentenças distratoras e a trinta sentenças distribuídas em dez paradigmas, sendo que cada um deles continha três sentenças: uma com oração complemento, uma com o verbo convencer e uma com oração adjunta, todas as três com extração de elemento wh-⁵ da posição de sujeito da oração encaixada finita. A Tabela 1 seguinte apresenta a distribuição das sentenças do estudo.

⁴ Testes de Percepção/Tarefas de Treinamento Perceptual da empresa Worken (<http://www.worken.com.br/tp>)

⁵ Tipo de sintagma interrogativo, neste estudo, classificamos em dois tipos: simples e complexo.

Tabela 1: Distribuição das sentenças.

Tipo de sentença	Sintagma interrogativo simples	Sintagma interrogativo complexo	Total
Familiarização	6	2	8
Complemento	5	5	10
Adjunto	6	4	10
Convencer	3	7	10
Distratora	3	5	8
TOTAL	23	23	46

O objetivo do teste é obter julgamentos para as sentenças com o verbo convencer e observar se eles se aproximam dos julgamentos para as sentenças com orações complemento ou das sentenças com orações adjuntas. As sentenças foram formuladas com o máximo de atenção para que questões pragmáticas não interferissem nos julgamentos.

2.2. Descrição das variáveis

As variáveis analisadas neste estudo estão descritas a seguir:

- Resposta de participante:
Sim
Não
- Resposta esperada para sentenças que não são do tipo convencer:
Sim
Não
- Ordem de apresentação das sentenças:
Valores: 1 a 46
- Tipo de sentença:
Familiarização

Complemento

Adjunto

Convencer

Distratora

- Tipo de sintagma interrogativo:
 - Simplex
 - Complexo
- Acionamento da tecla REPETIR:
 - Sim
 - Não

3. Exemplos

O objetivo destes exemplos é ajudar os leitores a entender as palavras específicas da área linguística e do relatório.

O João prometeu para Maria **que ia sair**.

O elemento "que ia sair" é complemento do verbo. O verbo "prometer" de dois objetos "a Maria" "que ia sair" para completar o significado da frase. Essa sentença é um exemplo da sentença do tipo *adjunto*.

O João viu a Maria **quando viajou**.

Nessa sentença, o elemento "quando viajou" é um adjunto. O verbo "ver" não precisa de mais esse objeto, isso é informação opcional sobre essa sentença. A sentença a seguir é um exemplo da sentença do tipo *complemento*

Que funcionário o João convenceu a Maria **que devia sair da sala**?

A sentença acima é uma sentença do teste que pertence o tipo *convencer*. Se o participante acha que a sentença é bem formada, implica que ele considera "que devia sair da sala" é um complemento da sentença, e se o participante acha que a sentença é mal formada, significa que ele considera "que devia sair da sala" é um adjunto da sentença.

4. Análise descritiva

Com o intuito de verificar o objetivo do estudo, foi feita uma análise descritiva dos dados coletados, baseadas em gráficos de barras de números absolutos, de proporções, gráficos de dispersão e box-plots. Para entender melhor essas técnicas usadas, se recomenda o livro de Bussab e Morettin (2013).

4.1. Sentenças dos tipos *familiarização*, *distratora* e *convencer*

Primeiramente, podemos observar no Gráfico B.1 e na Tabela A.1 o comportamento global do sujeito nulo nas sentenças com o verbo "convencer". A maioria das respostas dos participantes foram "Sim" nesse tipo de sentença, ou seja, a maioria deles consideraram as sentenças do tipo *convencer* como sentença que tem função de complementos.

Agora, vamos verificar se alguma sentença teve comportamento atípico, ou seja, se existe alguma sentença que muitos participantes acertaram ou erraram dentro do tipo da mesma. O Gráfico B.2 e a Tabela A.2 nos mostram que as sentenças do tipo *familiarização* não tiveram comportamento atípico, pois não existe *outliers* e para as sentenças do tipo *distratora*, podemos tirar a mesma conclusão, através do Gráfico B.3 e da Tabela A.3.

Analisando a Tabela A.4 e o Gráfico B.4 podemos perceber que a maioria dos participantes acertaram as sentenças do tipo *familiarização* e do tipo *distratora*, o que é um indício que os participantes entenderam o teste e prestaram atenção. Porém, nesse contexto, os participantes responderam mais "Sim" nas sentenças do tipo *convencer*, e pela prévia da pesquisadora a resposta correta para as sentenças do tipo *convencer* deve ser "Não".

A partir da Tabela A.5 e do Gráfico B.5, observamos que no gráfico de dispersão não existe tendência e podemos calcular o coeficiente de correlação linear entre as variáveis do eixo x e y, que é 0,24, o que implica que não há correlação entre os resultados individuais corretos das sentenças *familiarização* e os resultados individuais para estruturas com *convencer*.

No Gráfico B.6, que é gerado a partir da Tabela A.6, podemos obter as mesmas conclusões que no caso de *familiarização*, que não observamos tendência pelo gráfico de dispersão, e o valor de coeficiente de correlação fica muito pequeno (0,11), o que significa que os resultados individuais para estruturas com *convencer* não são afetados pelos resultados individuais corretos das sentenças *distratoras*.

4.2. Sentenças dos tipos *adjunto*, *complemento* e *convencer*

Os Gráficos B.7, B.8 e B.9 correspondentes às Tabelas A.7, A.8 e A.9, respectivamente, nos mostram que as sentenças dos tipos adjunto, complemento e convencer não têm comportamentos atípicos, pois não existem *outliers* no box-plot. Mas no Gráfico B.10 (Tabela A.10) que apresentam as sentenças S21, S31 e S32 (todas são sentenças do tipo convencer e do tipo de sintagma interrogativo complexo) são *outliers*, ou seja, elas têm comportamentos atípicos no quadro global.

Agora, excluindo as 3 sentenças, no Gráfico B.11 (Tabela A.11) observamos que comparando com o caso que tem todas as sentenças, a proporção dos participantes que responderam "Não" nas sentenças do tipo *convencer* aumentou. Fazendo também essa comparação segundo o comportamento individual no Gráfico B.12 (Tabela A.12), podemos concluir que para todos os participantes menos o sujeito 18, a proporção de respostas "Não" aumentou no caso sem *outliers*.

Através da Tabela A.13 e do Gráfico B.13 podemos ver que a proporção dos participantes que responderam "Sim" (que é resposta errada para as sentenças do tipo *adjunto*) nas sentenças é muito menor quando o tipo de sintagma interrogativo é *simples* e a proporção dos participantes que responderam "Não" nas sentenças é maior quando o tipo de sintagma interrogativo é *simples*, logo, podemos concluir que o tipo de sintagma interrogativo interfere nas respostas corretas ou erradas para as sentenças *adjuntas*.

Já no Gráfico B.14 e na Tabela A.14 não vemos o mesmo comportamento. A proporção dos participantes que responderam "Sim" (que é resposta correta para as sentenças do tipo *complemento*) nas sentenças são próximas para os sintagmas simples e complexos, isto é, para as sentenças do tipo *complemento*, o tipo de sintagma interrogativo não interfere nas respostas dos participantes.

Comparando o Gráfico B.15 com o Gráfico B.16 podemos tirar uma conclusão: as respostas dos participantes nas sentenças do tipo *convencer* segundo o tipo de sintagma interrogativo são mais parecidas com as respostas dos participantes nas sentenças do tipo *adjunto*. O Gráfico B.17 apresenta a proporção de resposta "Sim" e "Não" das sentenças do tipo de sintagma interrogativo simples. Podemos perceber que as respostas dos participantes nas sentenças do tipo *convencer* são mais parecidas com as respostas dos participantes nas sentenças do tipo *adjunto*. Já no Gráfico B.18, tiramos a conclusão diferente: para as respostas do tipo de sintagma interrogativo complexo, as respostas dos participantes nas sentenças do tipo *convencer* são mais parecidas com as respostas dos participantes nas sentenças do tipo *complemento*. Os dados que geraram estes gráficos estão nas Tabelas A.15 a A.18, respectivamente.

Observamos no Gráfico B.19 e na Tabela A.19, que há correlação positiva para essas duas variáveis de natureza quantitativa, pois quando um participante respondeu mais "Sim" nas sentenças do tipo *adjunto*, respondeu mais "Sim" nas sentenças do tipo *convencer* também.

A partir da Tabela A.20, do Gráfico B.20 e do Gráfico B.21 podemos ver que a ordem de apresentação das sentenças não afeta as respostas dos participantes, pois as proporções dos participantes que responderam corretamente são aproximadas para cada sentença. Já no Gráfico B.22 e na Tabela A.21, não vemos o mesmo comportamento, pois a proporção dos participantes que responderam "Não" nas sentenças do tipo *convencer* está diminuindo de acordo com o número da sentença apresentada.

Agora, veremos qual é a variável que está relacionada com o fator *repetir*.

O Gráfico B.23 e a Tabela A.22 mostram que nas sentenças do tipo *familiarização* há maior proporção de repetição, devido ao fato de que as oito primeiras sentenças são de tipo *familiarização*, e por isto, os participantes não estavam

acostumados ao teste. Depois, as sentenças do tipo *adjunto* possuem a segunda maior proporção de repetição comparado com as outras sentenças. No Gráfico B.24 e na Tabela A.23, as quantidades de repetição nas sentenças são parecidas para os dois tipos de sintagma interrogativo, logo, podemos concluir que o tipo de sintagma interrogativo não é o fator que causa repetição. Já no Gráfico B.25 e na Tabela A.24, podemos observar uma diminuição na frequência de vezes que os participantes pressionaram a tecla "Repetir" conforme muda-se a ordem de apresentação das sentenças, e concluimos que há correlação entre a ordem e a repetição.

Através da Tabela A.25 e do Gráfico B.26, observamos que os participantes acertaram um pouco mais quando pediram repetição das sentenças do tipo *adjunto* e *convencer*, mas para as sentenças do tipo *complemento*, eles erraram mais quando pediram repetição.

5. Análise inferencial

A fim de estudar os objetivos propostos por esse estudo, fui feita uma análise inferencial dos dados coletados, baseados em teste de independência, de correlação e de tendência. E depois, para extrapolar os resultados da amostra para toda a população de interesse, usamos modelo de regressão beta.

5.1. Testes de independências

Com a finalidade de testar a existência de dependência entre tipo de sintagma interrogativo e as respostas dos participantes nas sentenças do tipo adjunto, complemento, convencer e quantidade de repetição foram feitos testes exatos de Fisher⁶. Para realizar esses testes, medimos quantidade de respostas "Sim" ou "Não" dos participantes nas sentenças dos três tipos e quantidade de repetição segundo o tipo de sintagma interrogativo.

Podemos observar na Tabela C.1 que as respostas dos participantes nas sentenças do tipo complemento e quantidade de repetição não dependem do tipo de sintagma interrogativo, pois o valor-p fica grande (usamos 5% de nível de significância). As respostas dos participantes nas sentenças do tipo adjunto e convencer dependem do tipo de sintagma interrogativo, ou seja, para as sentenças do tipo adjunto e convencer, o tipo de sintagma interrogativo interfere nas respostas dos participantes.

5.2. Testes de correlação

Para fazer esses testes, para cada participante, medimos a quantidade de respostas "Sim" e "Não" das sentenças do tipo adjunto, complemento, convencer e a quantidade de acertos nas sentenças do tipo distratora e familiarização.

⁶ O teste é útil para dados categóricos, que resultam de classificação de objetos em duas maneiras diferentes; ele é usado para examinar a significância da associação (contingência) entre os dois tipos de classificação.

Na Tabela C.2, observamos que só a variável resposta "Sim" nas sentenças do tipo convencer e a variável resposta "Sim" nas sentenças do tipo adjunto possuem correlação significativa, pois o valor-p (0,03619) é menor que 5%. Esse resultado implica que quando um participante respondeu mais "Sim" nas sentenças do tipo adjunto, respondeu mais "Sim" nas sentenças do tipo convencer. Pelos coeficientes de correlação de Pearson⁷ e os valores-p, outras variáveis da tabela não mostram que existe correlação significativa.

5.3. Testes de tendência

As variáveis usadas são quantidade de sentenças que os participantes acertaram de cada sentença do tipo adjunto, complemento e convencer (consideramos que "Não" é resposta correta desse tipo de sentença) e a quantidade de repetição de cada sentença segundo a ordem de apresentação.

A partir da Tabela C.3, concluímos que a quantidade de repetição tem tendência, podemos falar que a quantidade de repetição apresenta tendência de decaimento conforme muda-se a ordem de apresentação. Na parte descritiva, concluímos que a ordem afeta as respostas do tipo convencer, mas pelo resultado de teste, obtemos conclusão oposta.

5.4. Modelo de regressão beta

Mesmo que o modelo de regressão normal linear seja bastante utilizado em análises empíricas, o modelo de regressão beta é mais adequado para este estudo cujo objetivo é estudar a relação entre variável resposta, que assume continuamente valores no intervalo (0, 1), e outras variáveis que afetam seu comportamento através de uma estrutura de regressão, pois a distribuição beta é bastante flexível para modelar proporções.

Para ajustar o modelo, usamos as seguintes variáveis:

⁷ Vamos definir o coeficiente de correlação de Pearson por R.

R = 1 significa uma correlação perfeita positiva entre as duas variáveis.

R = -1 significa uma correlação negativa perfeita entre as duas variáveis.

R = 0 significa que as duas variáveis não dependem linearmente uma da outra.

- Proporção de acerto para cada sentença
- Tipo de sentença (Adjunto, Complemento, Convencer, Distratora, Familiarização)
- Tipo de sintagma interrogativo (Complexo, Simples)
- Proporção de "repetir" para cada sentença
- Ordem da sentença (1 a 46)

O modelo utilizado pode ser escrito como:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + X^t\beta,$$

em que,

$\frac{p_i}{1-p_i}$ é a razão entre a proporção dos participantes acertaram e a proporção dos participantes erraram a i -ésima sentença;

X^t é o vetor (X_1, \dots, X_k) , sendo os elementos são os valores das variáveis explicativas;

β é o vetor dos coeficientes referentes às variáveis.

Inicialmente foi ajustado um modelo com todas as variáveis explicativas (Tipo de sentença, Tipo de sintagma interrogativo, Proporção de repetir para cada sentença e Ordem da sentença). Conforme a Tabela C.4, podemos observar na última coluna dos valores-p que Ordem, Tipo de sintagma interrogativo e Repetição não são variáveis significativas a um nível de 5%.

A partir dos resultados obtidos no ajuste do modelo completo, foi retirada a variável com maior valor-p, ajustando-se um novo modelo sem essa variável. Esse procedimento foi repetido até permanecerem no modelo somente variáveis cujos coeficientes foram significativos. Desta forma, os parâmetros estimados do modelo final estão na Tabela C.5.

Para comprovar o bom ajuste do modelo final, foi realizada análise de diagnóstico para os resíduos que pode ser vista através dos Gráficos D.1 a D.4. Podemos observar que a observação 30 (uma sentença do tipo familiarização com proporção de acerto 1) é considerada influente em relação aos demais.

Os parâmetros estimados do modelo sem a observação 30 estão na Tabela C.6. Comparado com o modelo final dos dados completos, familiarização torna-se um parâmetro não significativo. De acordo com os coeficientes da Tabela C.6, estimou-se que:

- A proporção que os participantes acertaram é aproximadamente $e^{1,38883} \approx 4,01$ vezes a proporção que os participantes erraram uma sentença do tipo *complemento*.
- A proporção que os participantes acertaram uma sentença do tipo *complemento* é $\frac{e^{1,48939}}{1 + e^{1,48939}} \approx 0,8$.
- A proporção que os participantes acertaram uma sentença do tipo *convencer* é $\frac{e^{1,48939-2,03312}}{1 + e^{1,48939-2,03312}} \approx 0,37$.
- A proporção que os participantes acertaram uma sentença do tipo *adjunto* é $\frac{e^{1,48939-0,77804}}{1 + e^{1,48939-0,77804}} \approx 0,66$.

Os Gráficos D.5 a D.8 estão indicando que o modelo está bem ajustado.

Conclusão

Com o resultado da análise descritiva e da análise inferencial, podemos concluir que o tipo de sintagma interrogativo interfere nas respostas dos participantes nas sentenças do tipo *adjunto* e *convencer*. A estrutura das respostas das sentenças do tipo *convencer* parece mais com a do tipo *adjunto* nas sentenças do tipo de sintagma interrogativo simples e parece mais com a do tipo *complemento* nas sentenças do tipo de sintagma interrogativo complexo. Os participantes que responderam mais "Sim" nas sentenças do tipo *adjunto*, responderam mais "Sim" nas sentenças do tipo *convencer* também. A ordem não interfere nas respostas dos participantes nas sentenças do tipo *adjunto*, *complemento* e *convencer*, porém, a quantidade de *repetir* diminui, conforme muda-se a ordem.

Para o modelo dos dados reduzidos, a proporção de acerto das sentenças do tipo *complemento* é igual com a proporção de acerto das sentenças do tipo *familiarização* e do tipo *distratora*. Comparado com as sentenças do tipo *adjunto* e *convencer*, a proporção de acerto das sentenças do tipo *complemento* possui a proporção maior.

Anexo 1 - As sentenças do teste

Sigla	Sentenças
S1	Em quem a Maria falou para Pedro que ela ia volta?
S2	Que animal a Maria viu quando foi ao zoológico com o Pedro?
S3	Que político o repórter que entrevistou foi despedido?
S4	Quem o João disse pro Pedro que a Maria visitou no sábado?
S5	O que o João convidou a Maria pra comer o bolo?
S6	Quem a Maria acha que o Pedro devia contratar?
S7	Quem o João disse pra Susana que o Pedro viu a Maria?
S8	Em quem a Maria ficou brava com o Pedro depois que ele votou?
S9	Quem a Maria contou pro Pedro que vem pra festa de amanhã?
S10	Que casa a Maria contou pro Pedro que foi vendido?
S11	Quem a Maria terminou com o Pedro depois que ela conheceu na festa?
S12	Quem o João convenceu o Pedro que ganhou na loteria?
S13	Que político a Maria reclamou pro partido depois que fez campanha na televisão?
S14	Que inspetor a Maria comentou com o João que era bom no que fazia?
S15	Que parente o João falou pro Pedro que a Maria visitou?
S16	Quem o João convenceu a Maria que vem pra festa amanhã?
S17	Que político o João falou pro Pedro que não cumpriu o mandato?
S18	Que funcionário o João convenceu a Maria que precisava comprar um carro novo?
S19	Que prato a Maria convidou o Pedro depois que cozinhou?
S20	Que diretor a Maria comentou com o Pedro que faltou na reunião?
S21	Que funcionário o João convenceu a Maria que devia sair da sala?
S22	Quem o vendedor informou a Maria que o produto chegou?
S23	Que aluno a Maria falou pra diretora que fez uma boa apresentação?
S24	Quem o João convenceu a Maria que vai viajar pra Holanda?
S25	Quem a Maria disse que o João chamou para sair?

S26	Que inspetor o João promoveu o Pedro quando foi demitido?
S27	Que pessoa a Maria convenceu o Pedro que devia comprar a casa?
S28	Que diretor o João chamou a Maria depois que mandou o comunicado da reunião?
S29	Que menina a Maria disse que o Pedro beijou ontem?
S30	Quem o Pedro viu a Maria enquanto fazia compras?
S31	Que político a Maria convenceu o João que tem ficha suja?
S32	Que aluno a Maria convenceu a diretora que fez uma boa apresentação?
S33	Quem o Pedro ligou pra Maria quando saiu da sala?
S34	Quem o Pedro disse pra Maria que viajou pra Londres?
S35	Que menina a Maria disse que o Pedro beijou a Susana ontem?
S36	Quem a Maria disse pro Pedro que precisava comprar um carro novo?
S37	Quem a Maria ligou pro Pedro quando disse que ia comprar a casa?
S38	Quem o João disse pro Pedro que ganhou na loteria?
S39	Quem a Maria encontrou o João depois que viajou pra Nova York?
S40	Quem a Maria convidou o Pedro depois que confirmou que vem pra festa?
S41	Que inspetor o Pedro convenceu o João que ia visitar a obra?
S42	Quem a Maria falou pro Pedro que teve que sair da sala?
S43	Quem o João discutiu com o Pedro depois que ganhou na loteria?
S44	Que gerente a Maria convenceu o Pedro que ia faltar na reunião amanhã?
S45	Que cachorro o João disse que adotou?
S46	Que aluno a Maria conversou com a diretora depois que fez uma boa apresentação?

Apêndice A – Tabelas da Análise Descritiva

Tabela A.1 – Resposta de participantes nas sentenças do tipo *convencer*

TIPO DE SENTENÇA	RESPOSTA "SIM"	RESPOSTA "NÃO"	TOTAL
CONVENCER	119	71	190

Tabela A.2 - Quantidade de participantes que acertaram ou erraram em cada sentença do tipo *familiarização*

SENTENÇA	ACERTARAM	ERRARAM
S1	15	4
S2	19	0
S3	15	4
S4	13	6
S5	14	5
S6	18	1
S7	16	3
S8	12	7
MÉDIA	15	4

Tabela A.3 - Quantidade de participantes que acertaram ou erraram na cada sentença do tipo *distratora*

SENTENÇA	ACERTARAM	ERRARAM
S11	16	3
S15	18	1
S19	14	5
S22	11	8
S25	16	3
S29	18	1
S35	13	6
S45	18	1
MÉDIA	16	3

Tabela A.4 – Proporção da quantidade de respostas certas ou erradas nas sentenças segundo o seu tipo

TIPO DE SENTENÇA	RESPOSTA CERTA	RESPOSTA ERRADA	TOTAL
DISTRATORA	0,82	0,18	1
FAMILIARIZAÇÃO	0,80	0,20	1
CONVENCER	0,37	0,63	1

Tabela A.5 - Comportamento individual dos participantes nas sentenças do tipo *familiarização* e *convencer*

PARTICIPANTES	ACERTOU NO FAMILIARIZAÇÃO	RESPONDEU "NÃO" NO CONVENCER
SUJEITO1	7	1
SUJEITO2	4	1
SUJEITO3	6	3
SUJEITO4	6	7
SUJEITO5	8	8
SUJEITO6	4	6
SUJEITO7	8	0
SUJEITO8	7	6
SUJEITO9	6	6
SUJEITO10	5	2
SUJEITO11	7	6
SUJEITO12	7	0
SUJEITO13	7	3
SUJEITO14	7	0
SUJEITO15	7	2
SUJEITO16	6	0
SUJEITO17	4	5
SUJEITO18	8	7
SUJEITO19	8	8

Tabela A.6 - Comportamento individual nas sentenças do tipo *distratora* e *convencer*

PARTICIPANTES	ACERTOU NO DISTRATORA	RESPONDEU "NÃO" NO CONVENCER
SUJEITO1	6	1
SUJEITO2	6	1
SUJEITO3	6	3
SUJEITO4	5	7
SUJEITO5	7	8
SUJEITO6	8	6
SUJEITO7	7	0
SUJEITO8	8	6
SUJEITO9	6	6
SUJEITO10	7	2
SUJEITO11	7	6
SUJEITO12	7	0
SUJEITO13	8	3
SUJEITO14	6	0
SUJEITO15	7	2
SUJEITO16	4	0
SUJEITO17	4	5
SUJEITO18	8	7
SUJEITO19	7	8

Tabela A.7 - Quantidade de participantes que acertaram ou erraram em cada sentença do tipo *adjunto*

SENTENÇA	ACERTARAM	ERRARAM
S13	13	6
S26	12	7
S28	11	8
S30	16	3
S33	12	7
S37	14	5
S39	13	6
S40	15	4
S43	15	4
S46	8	11
MÉDIA	13	6

Tabela A.8 - Quantidade de participantes que acertaram ou erraram na cada sentença do tipo *complemento*

SENTENÇA	ACERTARAM	ERRARAM
S9	14	5
S10	12	7
S14	16	3
S17	18	1
S20	17	2
S23	17	2
S34	17	2
S36	18	1
S38	12	7
S43	15	4
MÉDIA	13	6

Tabela A.9 - Quantidade de participantes que acertaram ou erraram na cada sentença do tipo *convencer*

SENTENÇA	RESPONDERAM "NÃO"	RESPONDERAM "SIM"
S12	14	5
S16	8	11
S18	6	13
S21	4	15
S24	12	7
S27	8	11
S31	2	17
S32	3	16
S41	8	11
S44	6	13
MÉDIA	7	12

Tabela A.10 – Proporção de participantes que acertaram em cada sentença

SENTENÇA	PROPORÇÃO DE ACERTO	SENTENÇA	PROPORÇÃO DE ACERTO
S13	0,68	S10	0,63
S26	0,63	S14	0,84
S28	0,58	S17	0,95
S30	0,84	S20	0,89
S33	0,63	S23	0,89
S37	0,74	S34	0,89
S39	0,68	S36	0,95
S40	0,79	S38	0,63
S43	0,79	S42	0,79
S46	0,42	S45	0,95
S9	0,74	S12	0,74
S16	0,42	S22	0,58
S18	0,32	S25	0,84

S21	0,21	S29	0,95
S24	0,63	S35	0,68
S27	0,42	S1	0,79
S31	0,11	S2	1,00
S32	0,16	S3	0,79
S41	0,42	S4	0,68
S44	0,32	S5	0,74
S11	0,84	S6	0,95
S15	0,95	S7	0,84
S19	0,74	S8	0,63

Tabela A.11 - Proporção dos participantes que responderam “Não” nas sentenças do tipo *convencer* para caso diferente

CASO	RESPONDERAM "SIM"	RESPONDERAM "NÃO"
TODAS AS SENTENÇAS	0,63	0,37
TIRAR 3 SENTENÇAS	0,53	0,47

Tabela A.12 - Proporção do participante que respondeu “Não” nas sentenças do tipo *convencer* para caso diferente

PARTICIPANTES	TODAS AS SENTENÇAS	TIRAR 3 SENTENÇAS
SUJEITO1	0,14	0,10
SUJEITO2	0,14	0,10
SUJEITO3	0,43	0,30
SUJEITO4	0,71	0,70
SUJEITO5	1,00	0,80
SUJEITO6	0,86	0,60
SUJEITO7	0,00	0,00
SUJEITO8	0,86	0,60
SUJEITO9	0,86	0,60
SUJEITO10	0,29	0,20

SUJEITO11	0,71	0,60
SUJEITO12	0,00	0,00
SUJEITO13	0,43	0,30
SUJEITO14	0,00	0,00
SUJEITO15	0,29	0,20
SUJEITO16	0,00	0,00
SUJEITO17	0,71	0,50
SUJEITO18	0,57	0,70
SUJEITO19	0,86	0,80
MÉDIA	0,47	0,37

Tabela A.13 - Proporção dos participantes que responderam "Sim" ou "Não" nas sentenças do tipo *adjunto* segundo o tipo de sintagma interrogativo

TIPO DE SINTAGMA INTERROGATIVO	RESPONDERAM "SIM"	RESPONDERAM "NÃO"
SIMPLES	0,25	0,75
COMPLEXO	0,42	0,58
TOTAL	0,32	0,68

Tabela A.14 - Proporção dos participantes que responderam "Sim" ou "Não" nas sentenças do tipo *complemento* segundo o tipo de sintagma interrogativo

TIPO DE SINTAGMA INTERROGATIVO	RESPONDERAM "SIM"	RESPONDERAM "NÃO"
SIMPLES	0,80	0,20
COMPLEXO	0,84	0,16
TOTAL	0,82	0,18

Tabela A.15 - Proporção dos participantes que responderam "Não" nas sentenças do tipo *adjunto*, *complemento* e *convencer* segundo tipo de sintagma interrogativo

TIPO DE SENTENÇA	COMPLEXO	SIMPLES
ADJUNTO	0,58	0,75
COMPLEMENTO	0,16	0,20
CONVENCER	0,28	0,60

Tabela A.16 - Proporção dos participantes que responderam "Sim" nas sentenças do tipo *adjunto*, *complemento* e *convencer* segundo tipo de sintagma interrogativo

TIPO DE SENTENÇA	COMPLEXO	SIMPLES
ADJUNTO	0,42	0,25
COMPLEMENTO	0,84	0,80
CONVENCER	0,72	0,40

Tabela A.17 - Proporção dos participantes que responderam "Sim" ou "Não" nas sentenças do tipo *adjunto*, *complemento* e *convencer* do tipo de sintagma simples

TIPO DE SENTENÇA	PROPORÇÃO DE SIM	PROPORÇÃO DE NÃO
ADJUNTO	0,42	0,58
COMPLEMENTO	0,84	0,16
CONVENCER	0,72	0,28

Tabela A.18 - Proporção dos participantes que responderam "Sim" nas sentenças do tipo *adjunto*, *complemento* e *convencer* do tipo de sintagma complexo

TIPO DE SENTENÇA	PROPORÇÃO DE SIM	PROPORÇÃO DE NÃO
ADJUNTO	0,25	0,75
COMPLEMENTO	0,80	0,20
CONVENCER	0,40	0,60

Tabela A.19 - Quantidade de respostas "Sim" de cada participante nas sentenças do tipo *adjunto* e *convencer*

PARTICIPANTE	RESPONDEU "SIM" NO ADJUNTO	RESPONDEU "SIM" NO CONVENCER
SUJEITO1	7	9
SUJEITO2	9	9
SUJEITO3	2	7
SUJEITO4	4	3
SUJEITO5	0	2
SUJEITO6	1	4
SUJEITO7	2	10
SUJEITO8	1	4
SUJEITO9	1	4
SUJEITO10	3	8
SUJEITO11	4	4
SUJEITO12	0	10
SUJEITO13	0	7
SUJEITO14	9	10
SUJEITO15	2	8
SUJEITO16	8	10
SUJEITO17	5	5
SUJEITO18	1	3
SUJEITO19	2	2
MÉDIA	3	6

Tabela A.20 – Proporção de acerto dos participantes para cada sentença do tipo *adjunto e complemento*

ADJUNTO	PROPORÇÃO DE ACERTO	COMPLEMENTO	PROPORÇÃO DE ACERTO
S13	0,68	S9	0,74
S26	0,63	S10	0,63
S28	0,58	S14	0,84
S30	0,84	S17	0,95
S33	0,63	S20	0,89
S37	0,74	S23	0,89
S39	0,68	S34	0,89
S40	0,79	S36	0,95
S43	0,79	S38	0,63
S46	0,42	S42	0,79
MÉDIA	0,68		0,82

Tabela A.21 – Proporção de resposta "Não" dos participantes nas sentenças do tipo *convencer*

SENTENÇA	PROPORÇÃO DE "NÃO"	SENTENÇA	PROPORÇÃO DE "NÃO"
S12	0,74	S27	0,42
S16	0,42	S31	0,11
S18	0,32	S32	0,16
S21	0,21	S41	0,42
S24	0,63	S44	0,32
MÉDIA	0,37		

Tabela A.22 – Proporção de repetição dos participantes para as sentenças segundo o seu tipo

TIPO DE SENTENÇA	SEM REPETIÇÃO	COM REPETIÇÃO
ADJUNTO	0,64	0,36
COMPLEMENTO	0,74	0,26
CONVENCER	0,71	0,29
DISTRATORA	0,70	0,30
FAMILIARIZAÇÃO	0,52	0,48

Tabela A.23 – Quantidade de repetição dos participantes para as sentenças segundo o seu tipo de sintagma interrogativo

TIPO DE SINTAGMA INTERROGATIVO	SEM REPETIÇÃO	COM REPETIÇÃO	PROPORÇÃO DE REPETIR
COMPLEXO	297	140	0,32
SIMPLES	283	154	0,35

Tabela A.24 – Quantidade de repetição dos participantes para as sentenças segundo a sua ordem

ORDEM	SIM	NÃO	PROPORÇÃO	ORDEM	SIM	NÃO	PROPORÇÃO
S1	11	8	0,58	S24	4	15	0,21
S2	8	11	0,42	S25	3	16	0,16
S3	8	11	0,42	S26	9	10	0,47
S4	8	11	0,42	S27	9	10	0,47
S5	11	8	0,58	S28	9	10	0,47
S6	4	15	0,21	S29	4	15	0,21
S7	10	9	0,53	S30	4	15	0,21
S8	13	6	0,68	S31	3	16	0,16
S9	7	12	0,37	S32	3	16	0,16

S10	7	12	0,37	S33	6	13	0,32
S11	8	11	0,42	S34	2	17	0,11
S12	8	11	0,42	S35	10	9	0,53
S13	10	9	0,53	S36	7	12	0,37
S14	5	14	0,26	S37	7	12	0,37
S15	5	14	0,26	S38	5	14	0,26
S16	7	12	0,37	S39	7	12	0,37
S17	3	16	0,16	S40	6	13	0,32
S18	10	9	0,53	S41	1	18	0,05
S19	9	10	0,47	S42	5	14	0,26
S20	7	12	0,37	S43	6	13	0,32
S21	4	15	0,21	S44	7	12	0,37
S22	5	14	0,26	S45	2	17	0,11
S23	10	9	0,11	S46	5	14	0,26

Tabela A.25 – Proporção de acerto dos participantes nas sentenças do tipo *adjunto*, *complemento* e *convencer* segundo o efeito de repetir

TIPO DE SENTENÇA	SEM REPETIÇÃO	COM REPETIÇÃO
ADJUNTO	0,65	0,72
COMPLEMENTO	0,86	0,72
CONVENCER	0,62	0,64

Apêndice B – Gráficos da Análise Descritiva

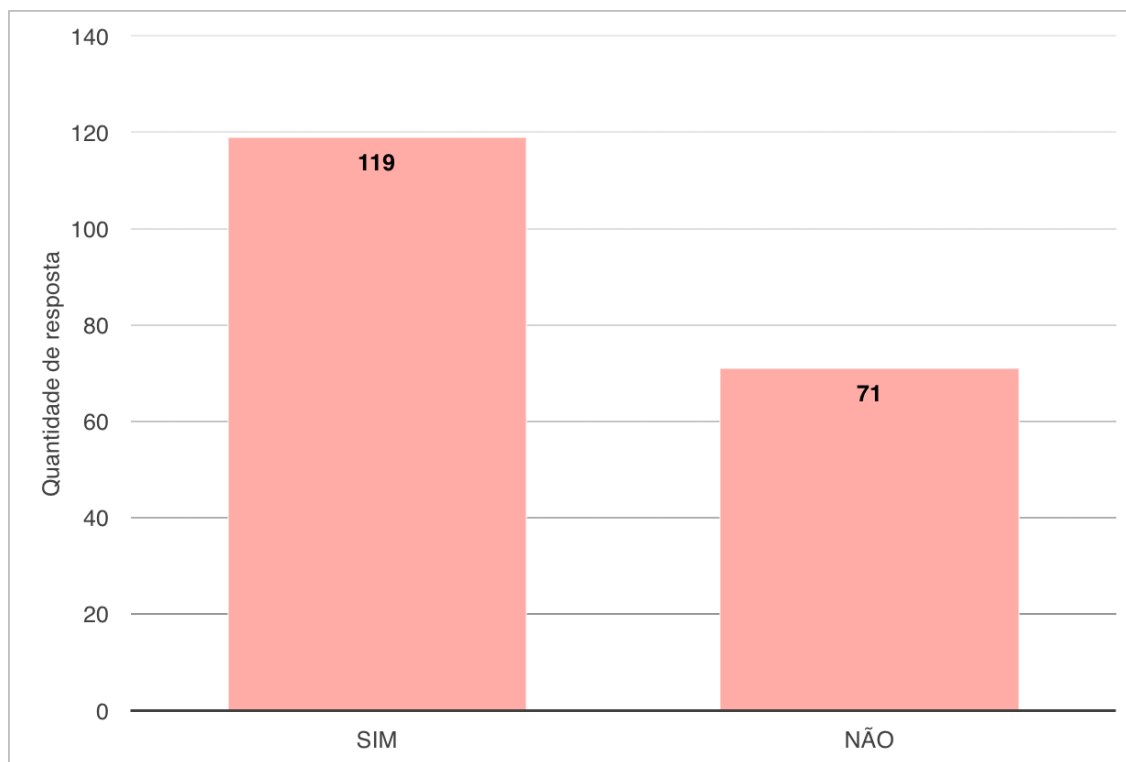
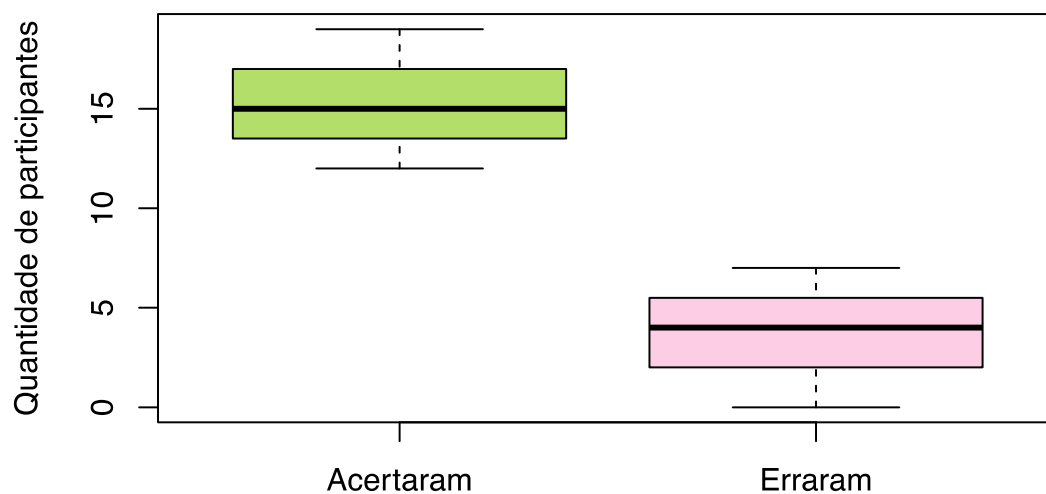
Gráfico B.1 - Quantidade de respostas "Sim" e "Não" do *tipo convencer***Gráfico B.2** - Box-plot de quantidade de participantes que acertaram ou erraram em cada sentença do *tipo familiarização*

Gráfico B.3 - Box-plot de quantidade de participantes que acertaram ou erraram em cada sentença do *tipo distratora*

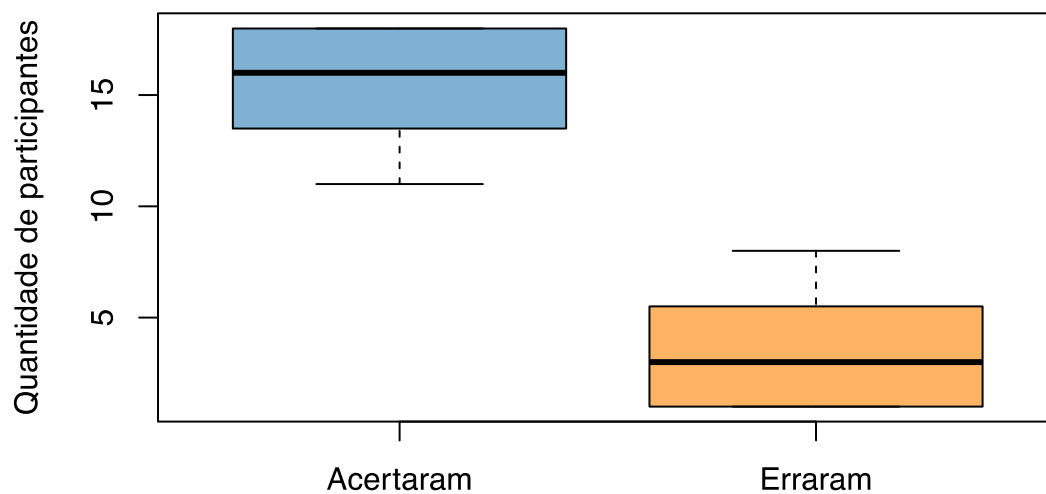


Gráfico B.4 - Proporção da quantidade de respostas certas ou erradas nas sentenças segundo o seu tipo

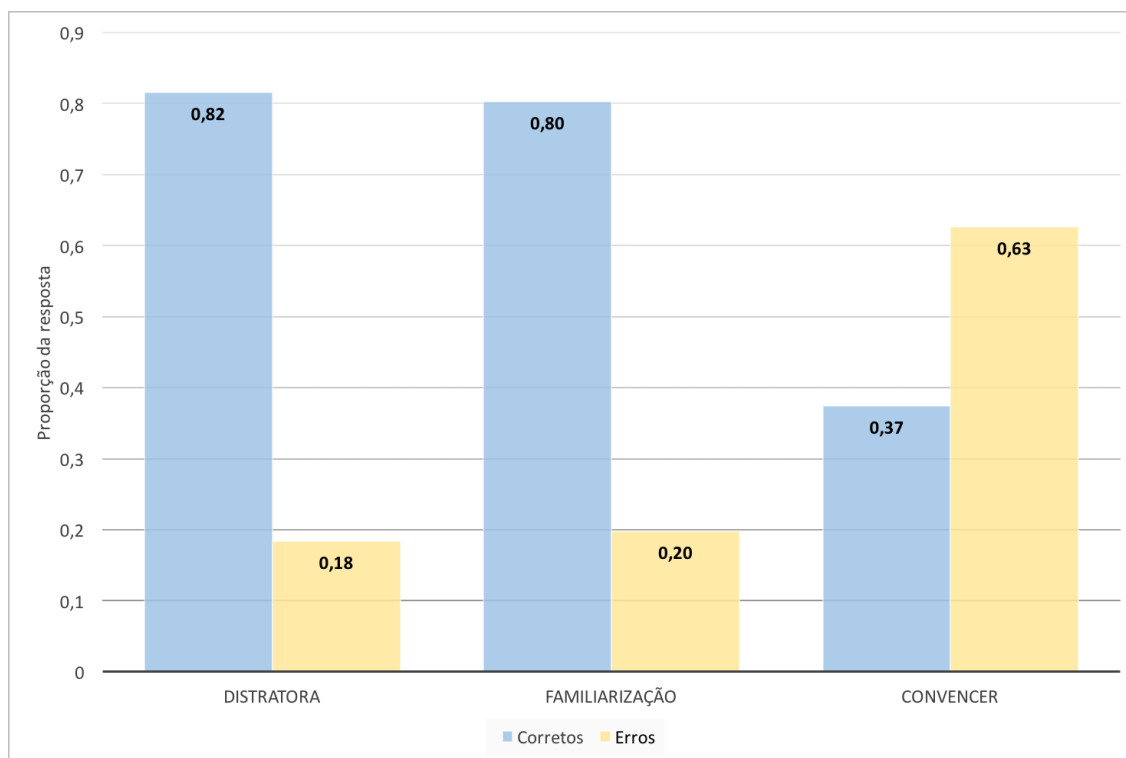


Gráfico B.5 – Comportamento individual dos participantes nas sentenças do tipo *familiarização e convencer*

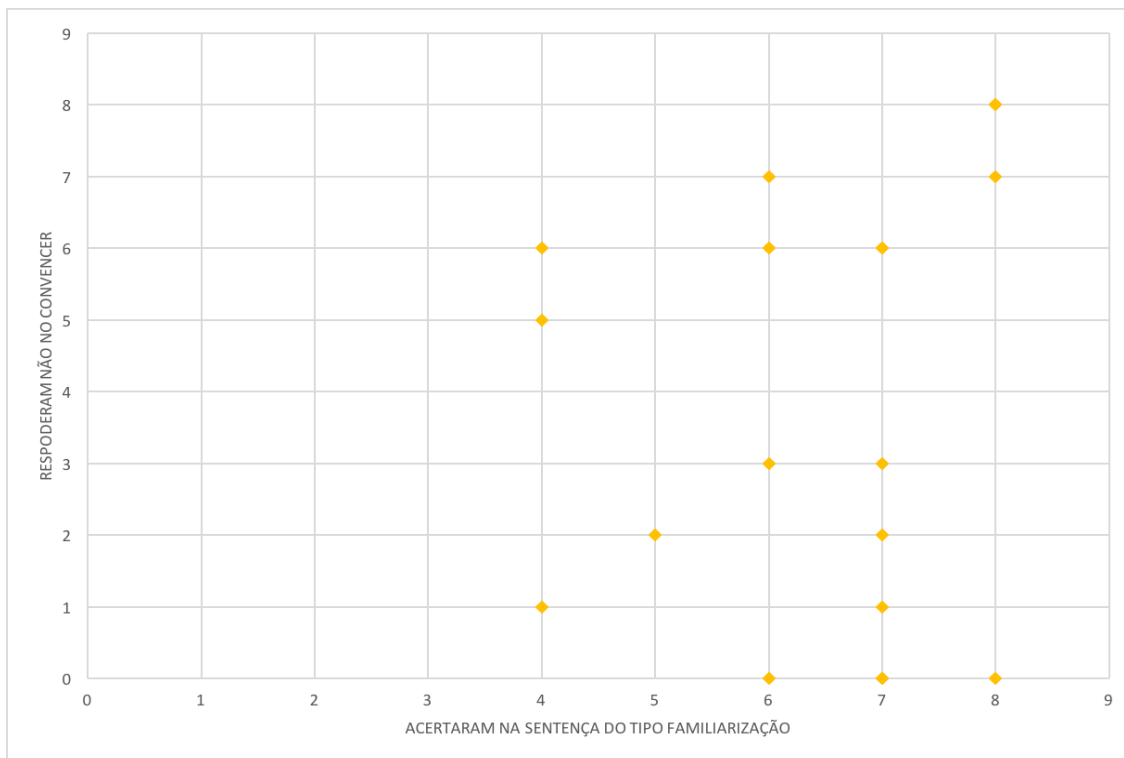


Gráfico B.6 – Comportamento individual dos participantes nas sentenças do tipo *distratora e convencer*

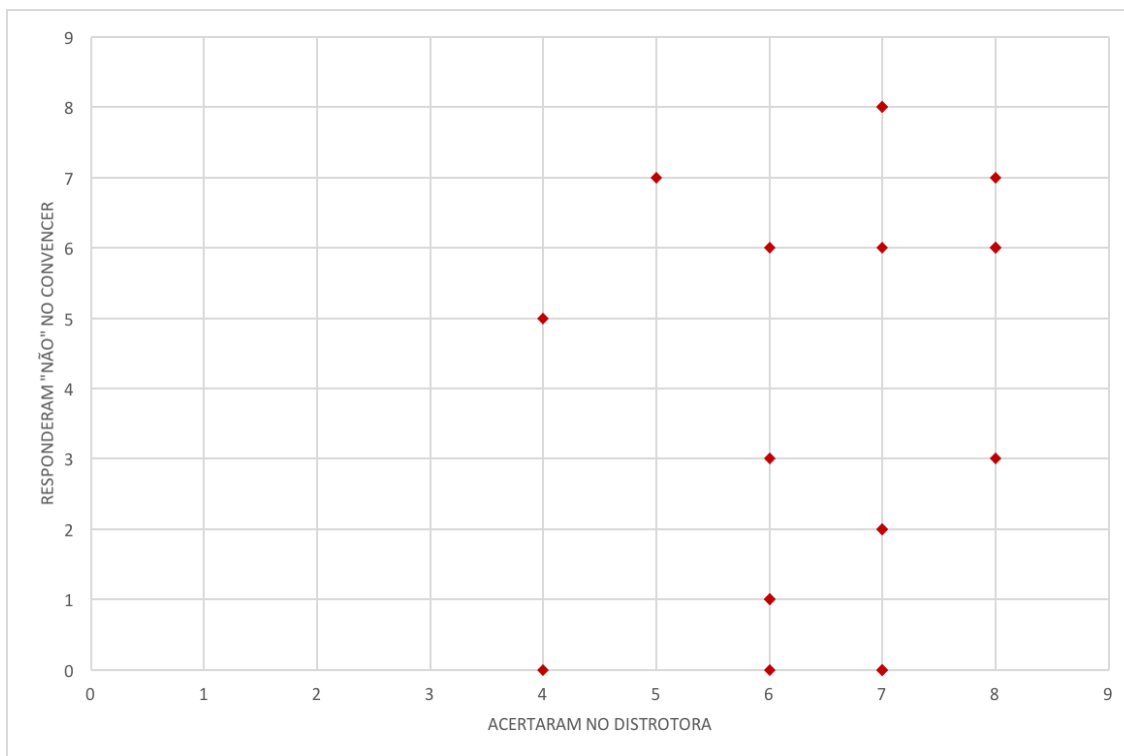


Gráfico B.7 - Box-plot da quantidade de participantes que acertaram ou erraram em cada sentença do tipo *adjunto*

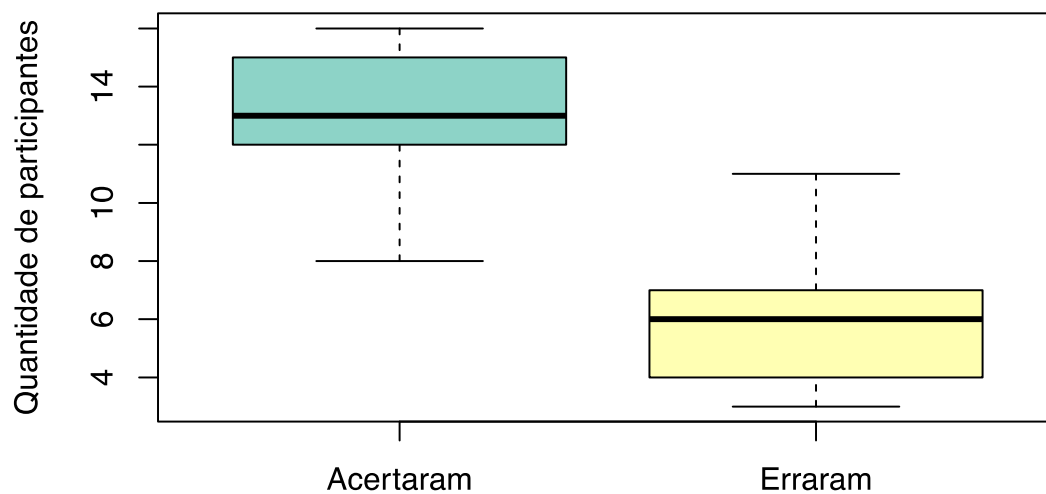


Gráfico B.8 - Box-plot da quantidade de participantes que acertaram ou erraram em cada sentença do tipo *complemento*

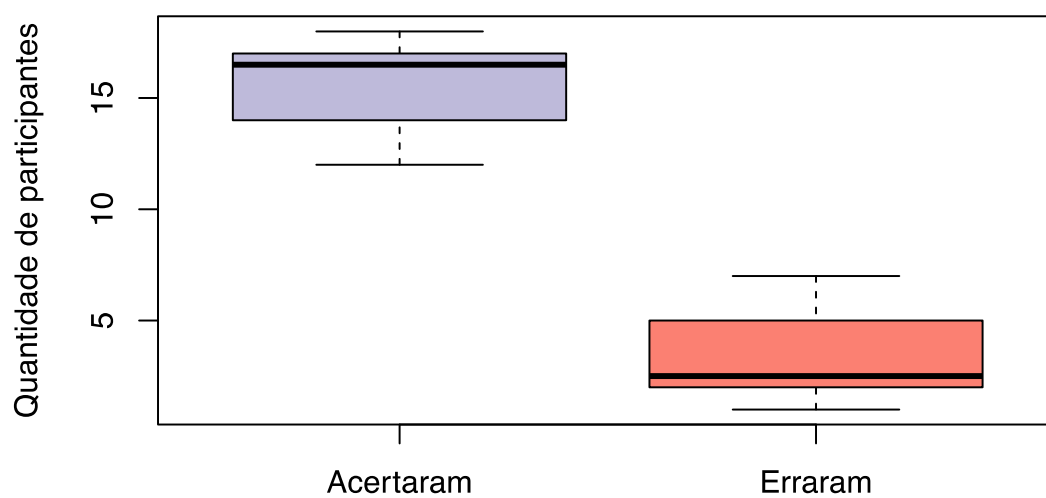


Gráfico B.9 - Box-plot da quantidade de participantes que acertaram ou erraram em cada sentença do tipo *convencer*

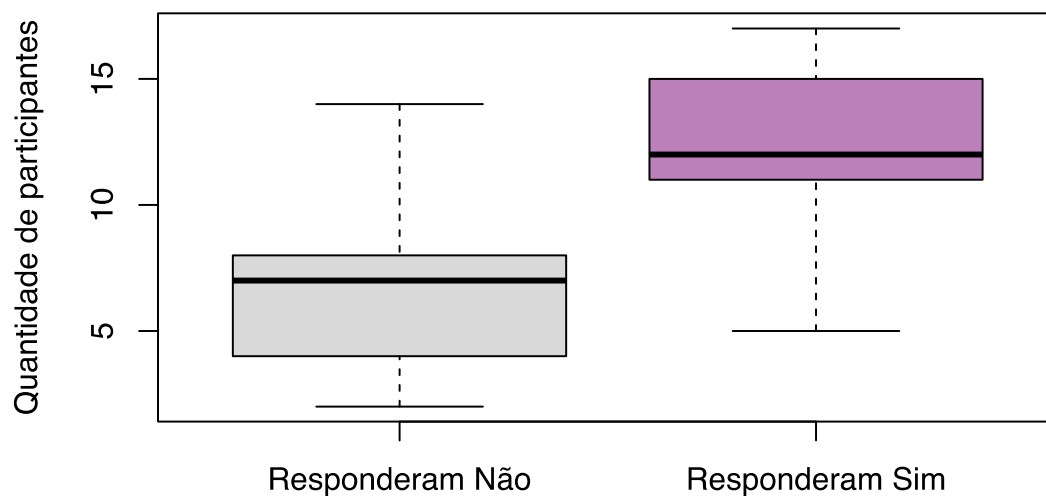


Gráfico B.10 - Box-plot da proporção de participantes que acertaram em cada sentença

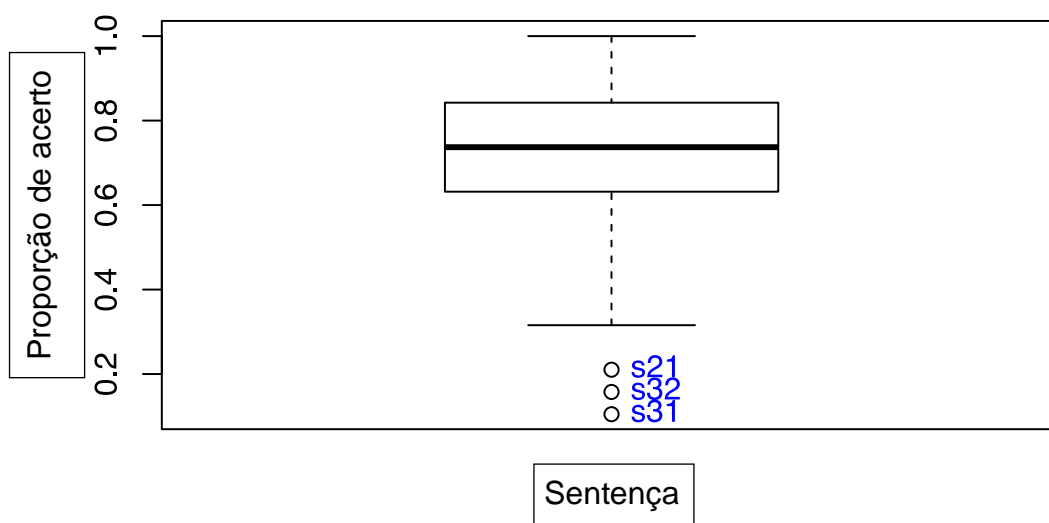


Gráfico B.11 - Proporção dos participantes que responderam “Não” nas sentenças do tipo *convencer* para caso diferente

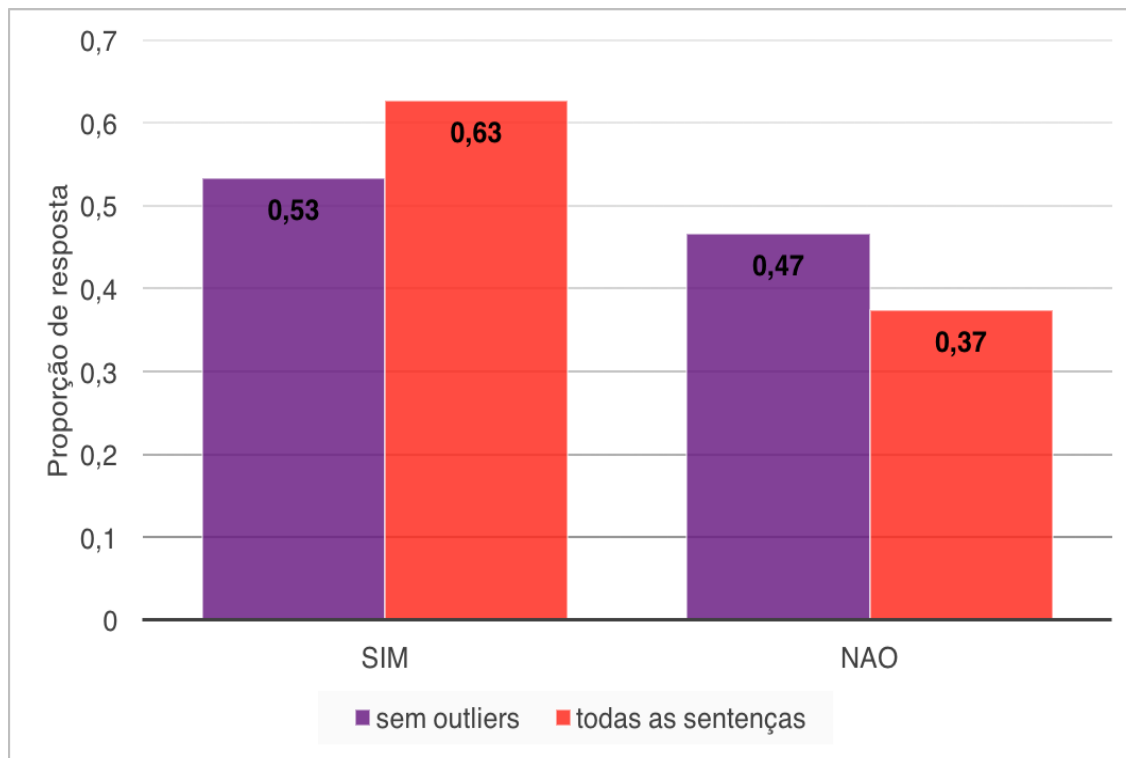


Gráfico B.12 - Proporção do participante que respondeu “Não” nas sentenças do tipo *convencer* para caso diferente

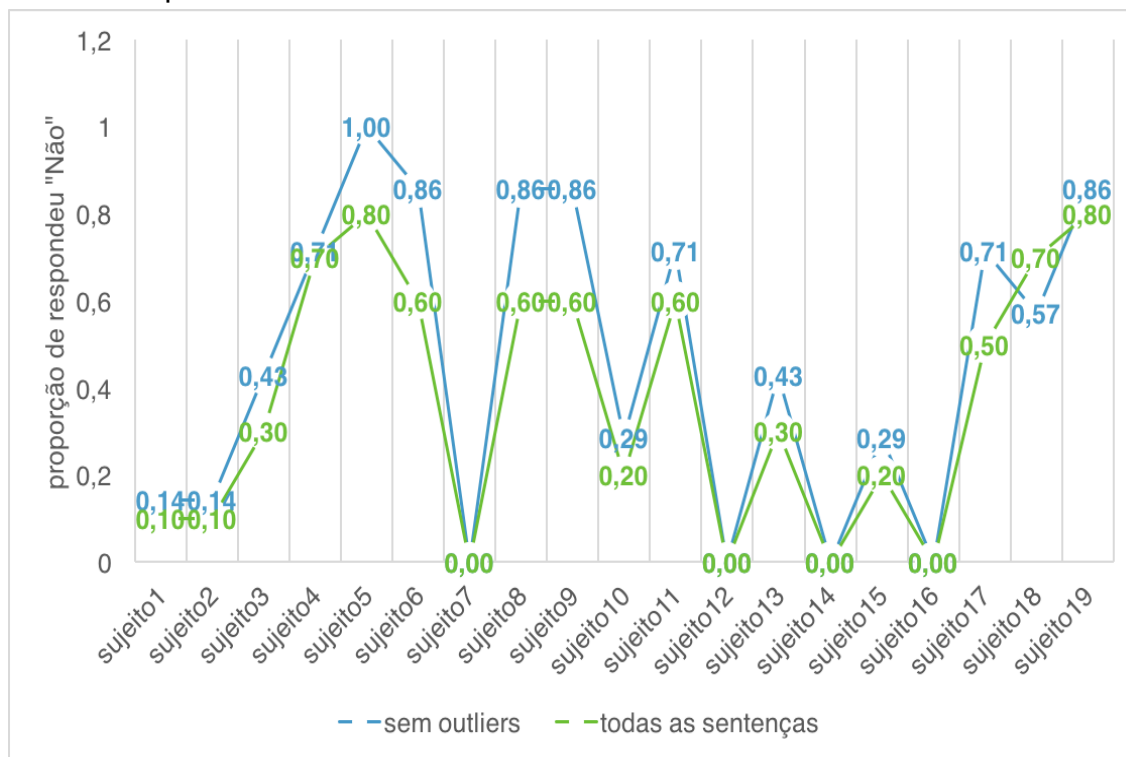


Gráfico B.13 - Proporção dos participantes que responderam "Sim" ou "Não" nas sentenças do tipo *adjunto* segundo o tipo de sintagma interrogativo

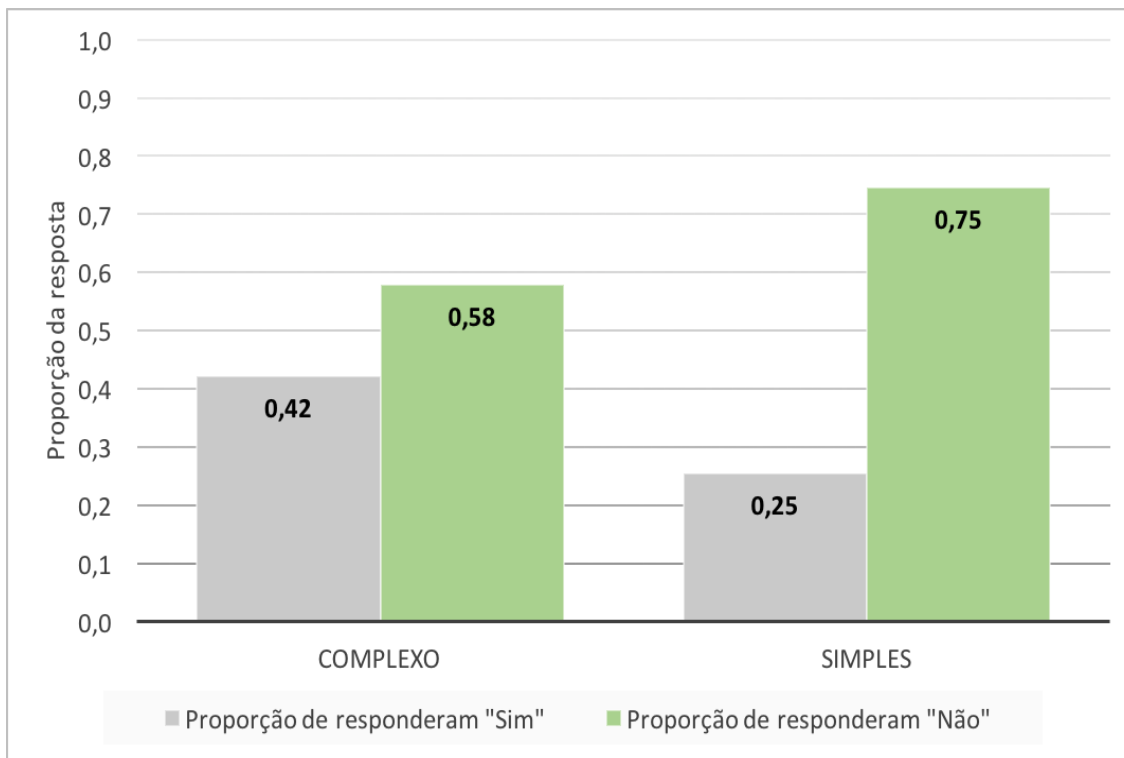


Gráfico B.14 - Proporção dos participantes que responderam "Sim" ou "Não" nas sentenças do tipo *complemento* segundo o tipo de sintagma interrogativo

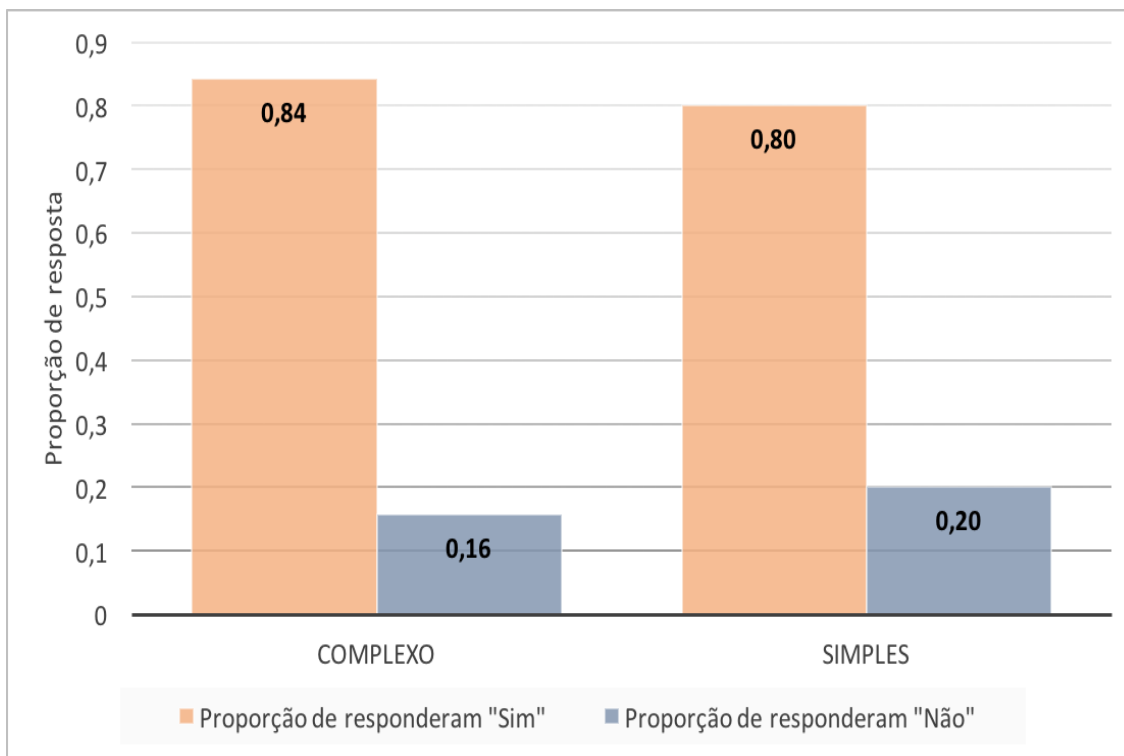


Gráfico B.15 - Proporção dos participantes que responderam "Não" nas sentenças do tipo *adjunto*, *complemento* e *convencer* segundo tipo de sintagma interrogativo

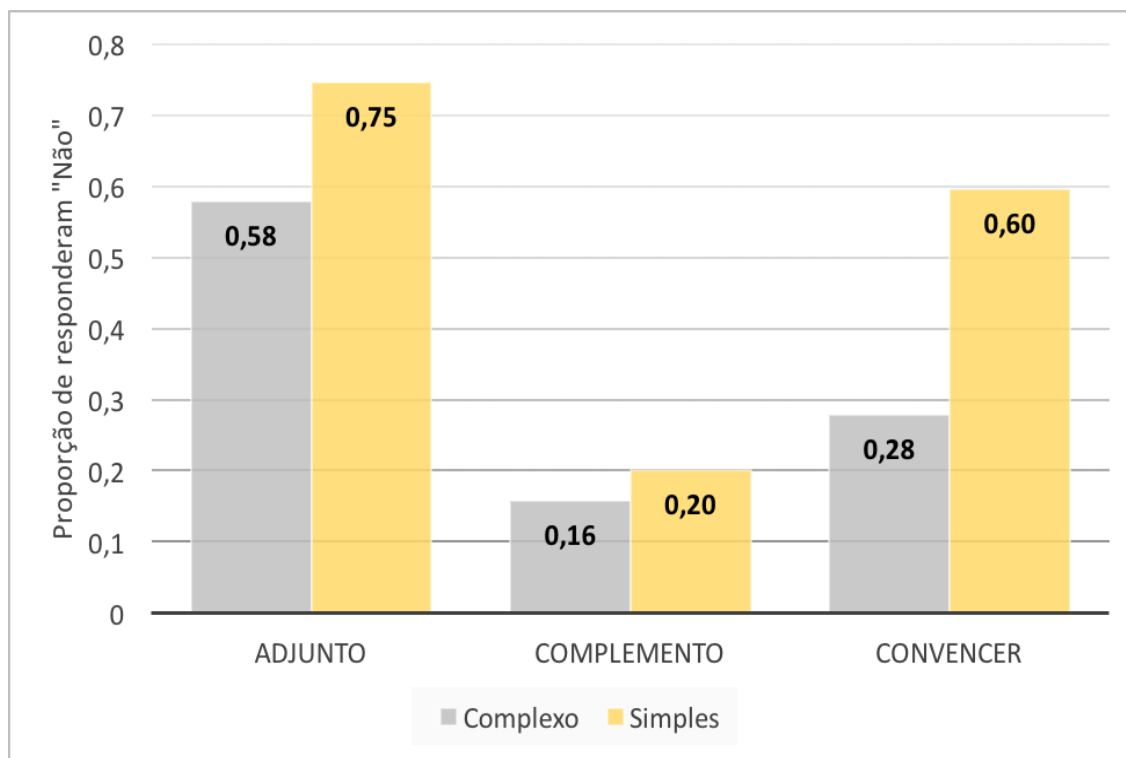


Gráfico B.16 - Proporção dos participantes que responderam "Sim" nas sentenças do tipo *adjunto*, *complemento* e *convencer* segundo tipo de sintagma interrogativo

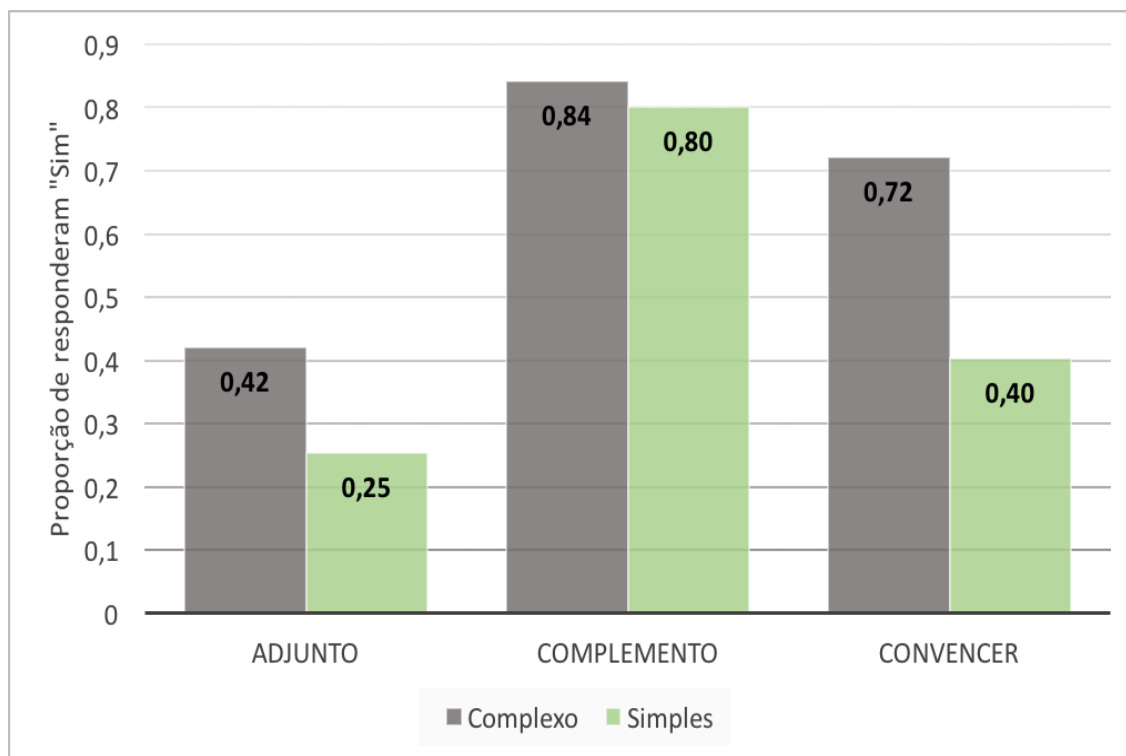


Gráfico B.17 - Proporção dos participantes que responderam "Sim" ou "Não" nas sentenças do tipo *adjunto*, *complemento* e *convencer* do tipo de sintagma simples

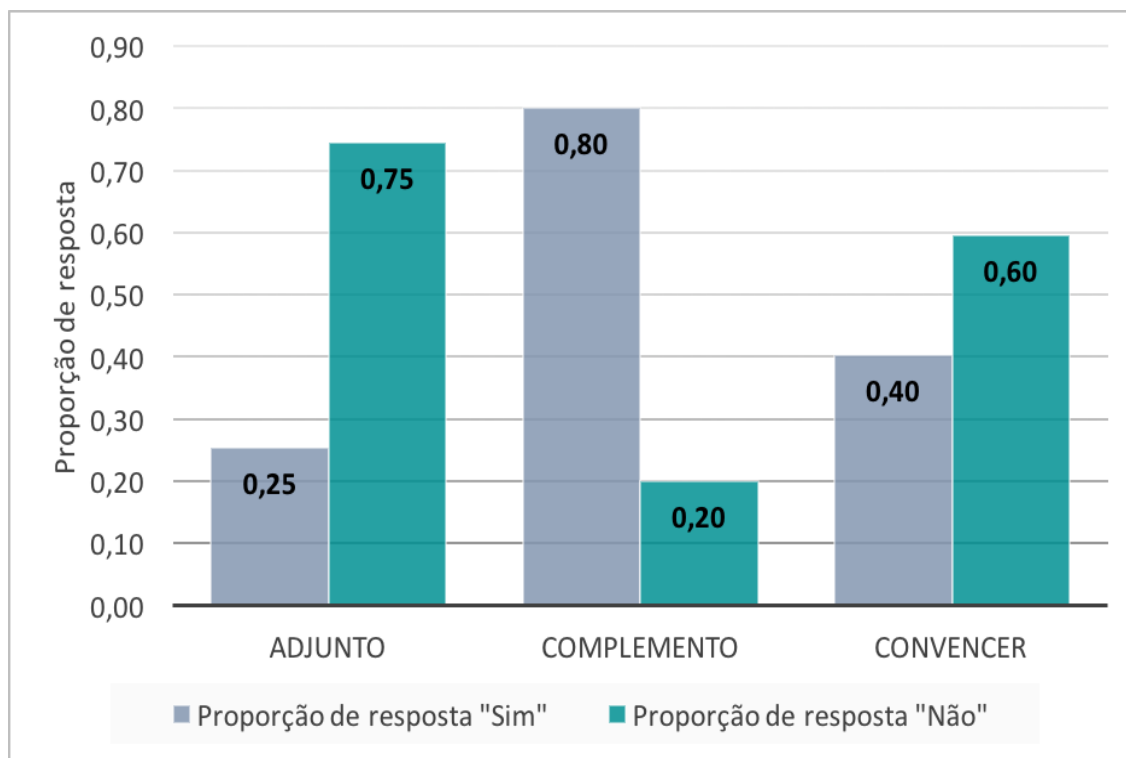


Gráfico B.18 - Proporção dos participantes que responderam "Sim" ou "Não" nas sentenças do tipo *adjunto*, *complemento* e *convencer* do tipo de sintagma complexo

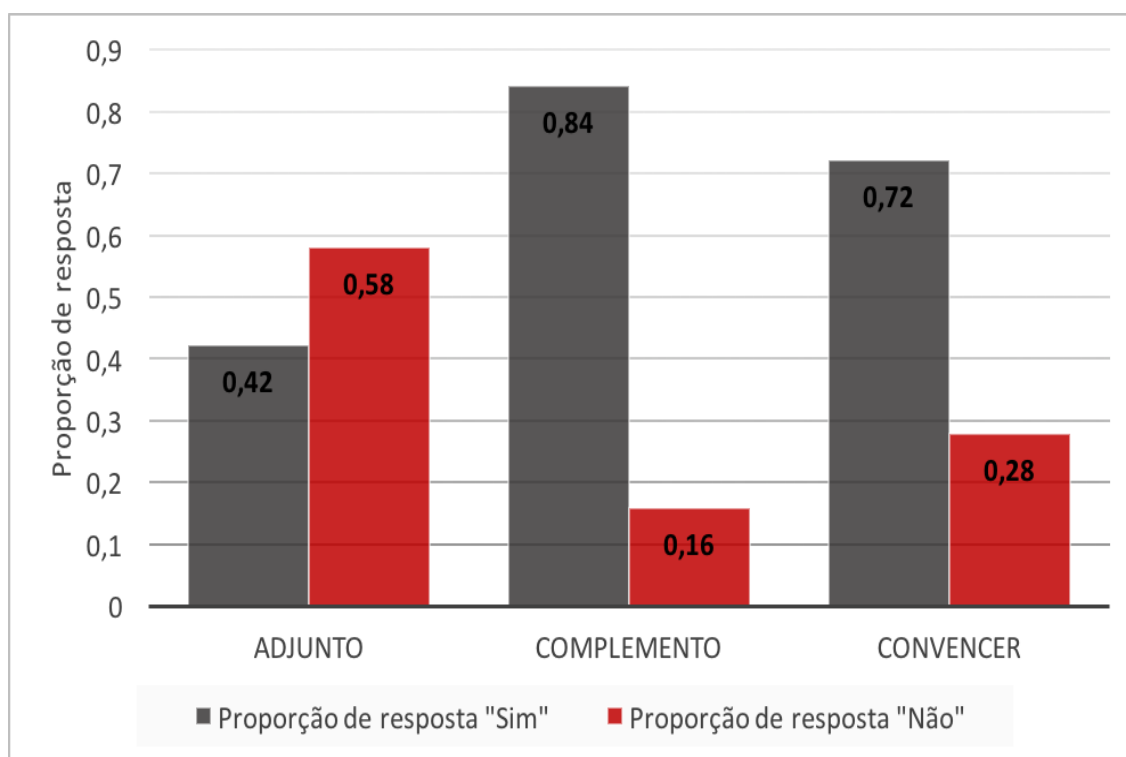


Gráfico B.19 - Quantidade de respostas "Sim" de cada participante nas sentenças do tipo *adjunto* e *convencer*

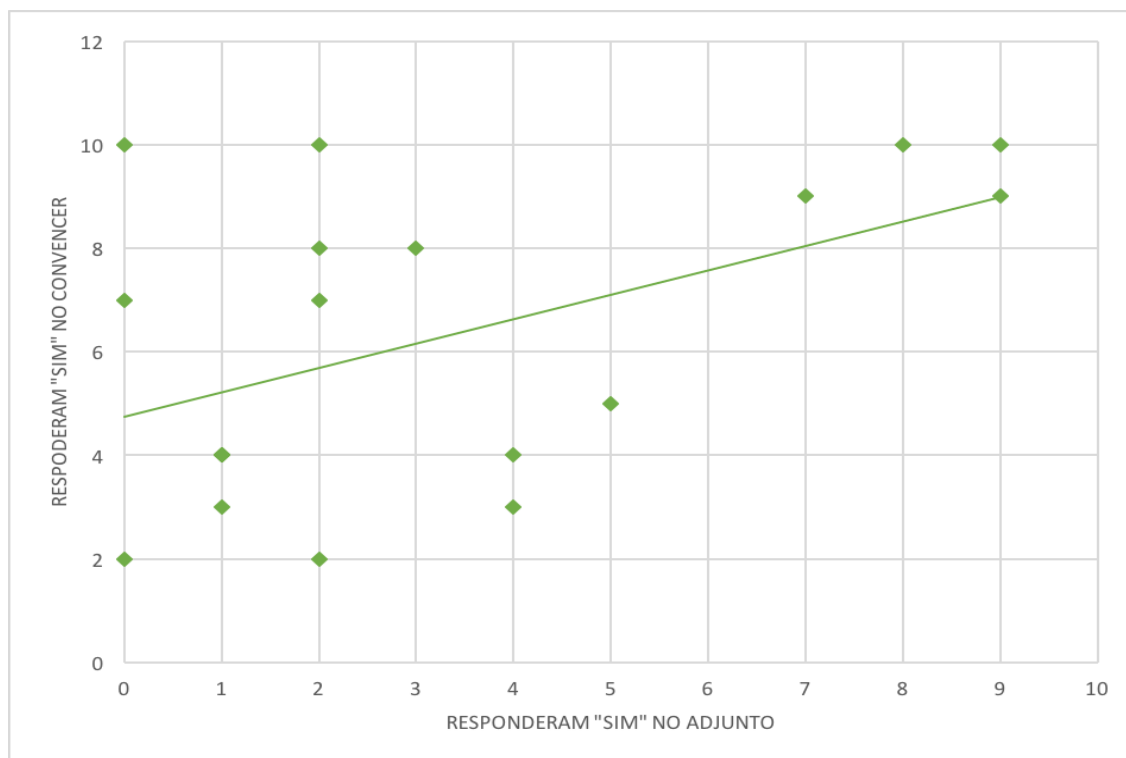


Gráfico B.20 - Proporção de acerto dos participantes para cada sentença do tipo *adjunto*

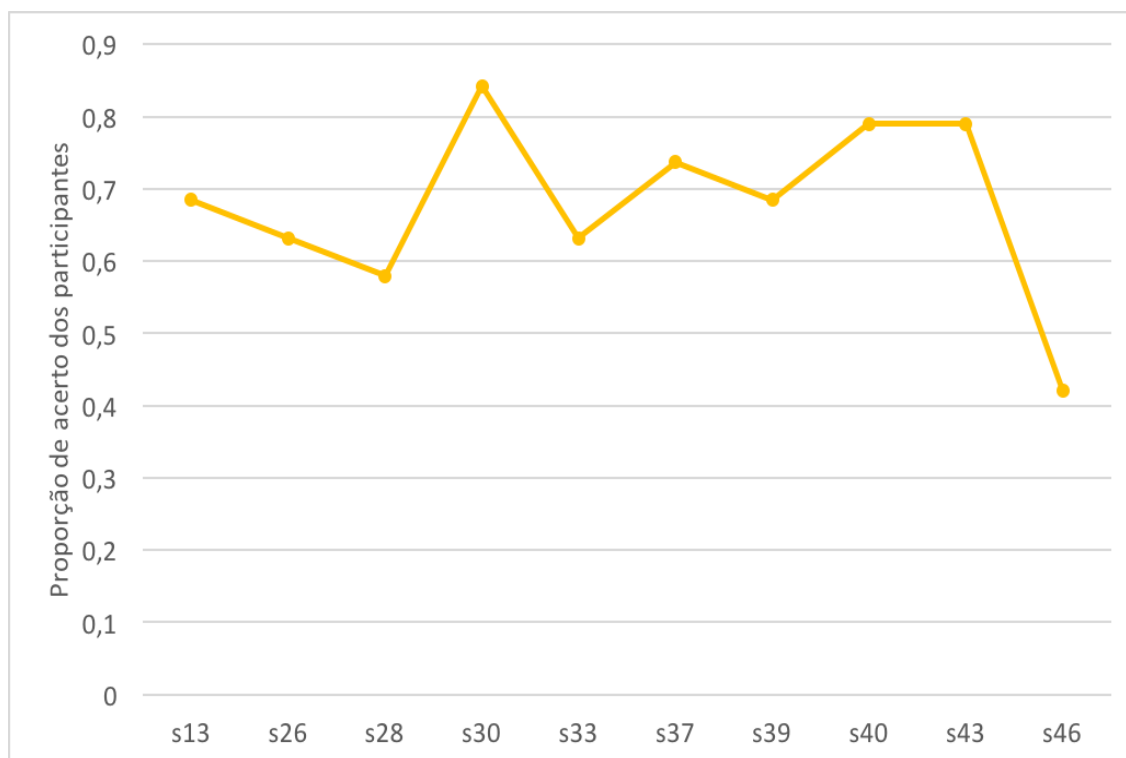


Gráfico B.21 - Proporção de acerto dos participantes para cada sentença do tipo *complemento*

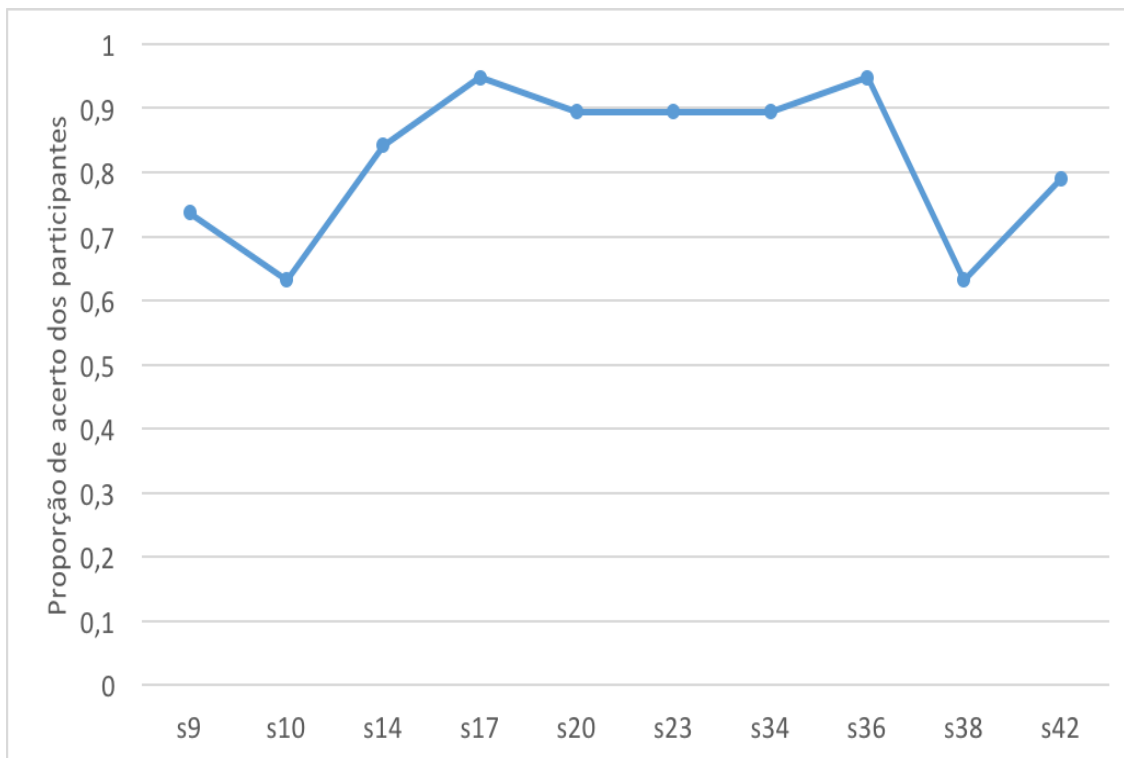


Gráfico B.22 - Proporção de resposta "Não" dos participantes nas sentenças do tipo *convencer*

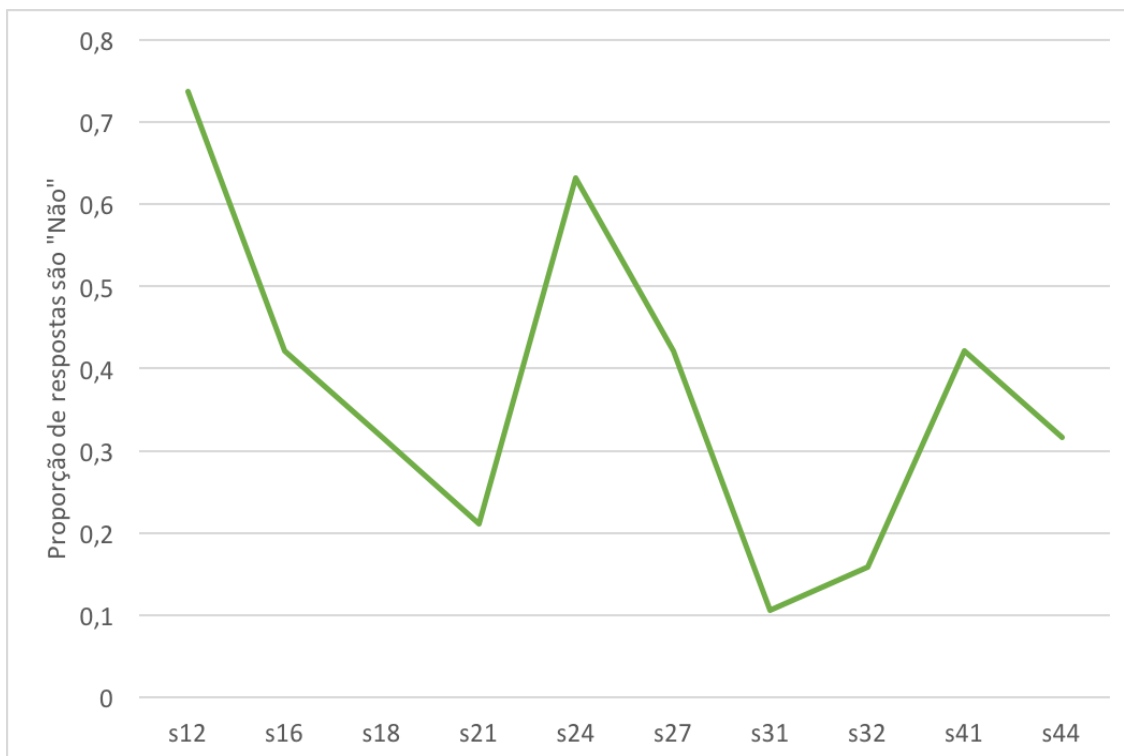


Gráfico B.23 – Proporção de repetição dos participantes para as sentenças segundo o seu tipo

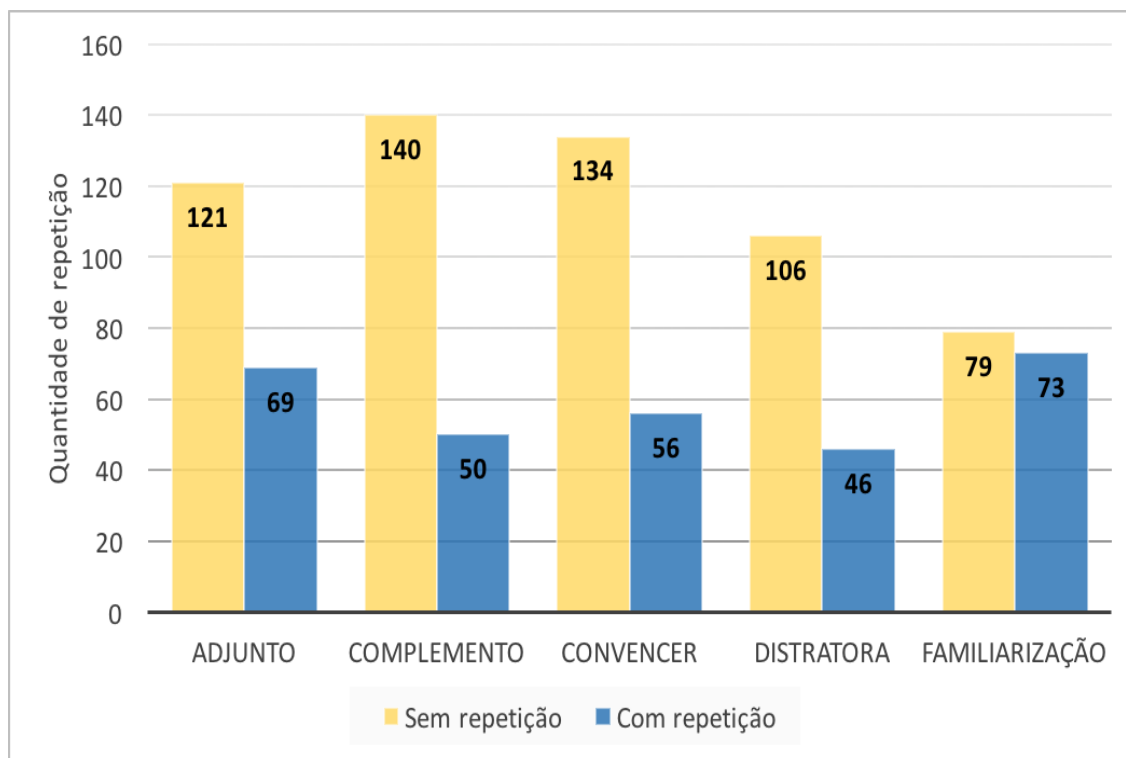


Gráfico B.24 – Quantidade de repetição dos participantes para as sentenças segundo o seu tipo de sintagma interrogativo

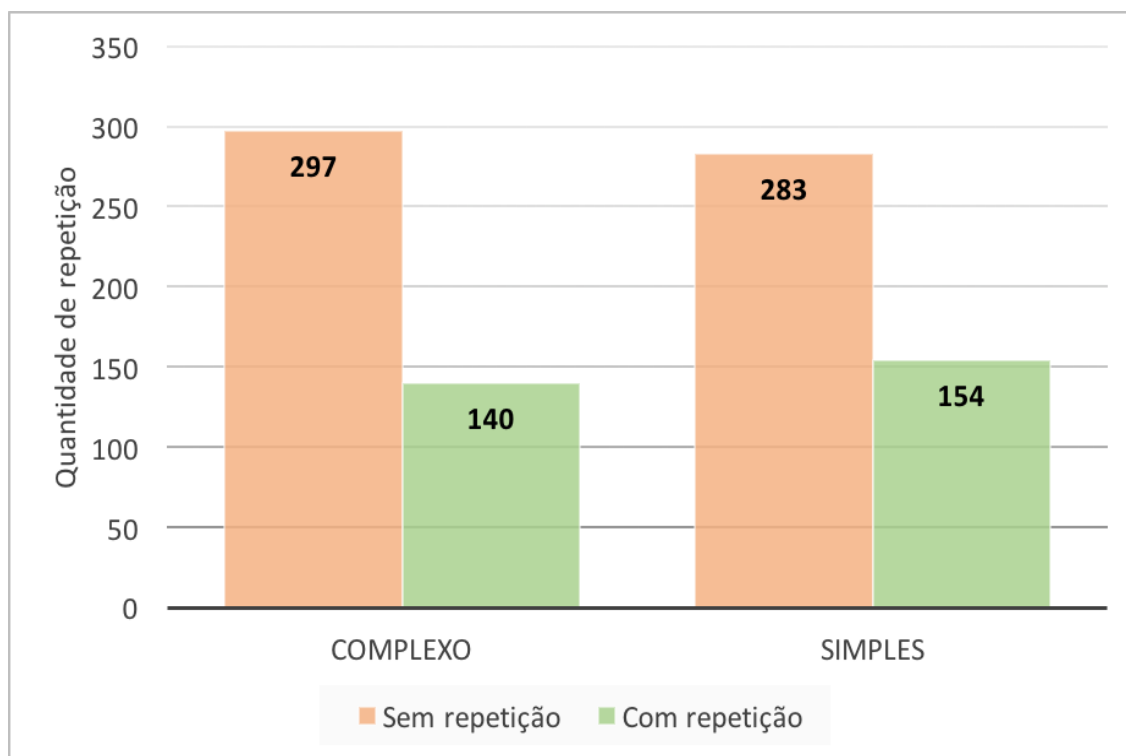


Gráfico B.25 - Quantidade de repetição dos participantes para as sentenças segundo a sua ordem

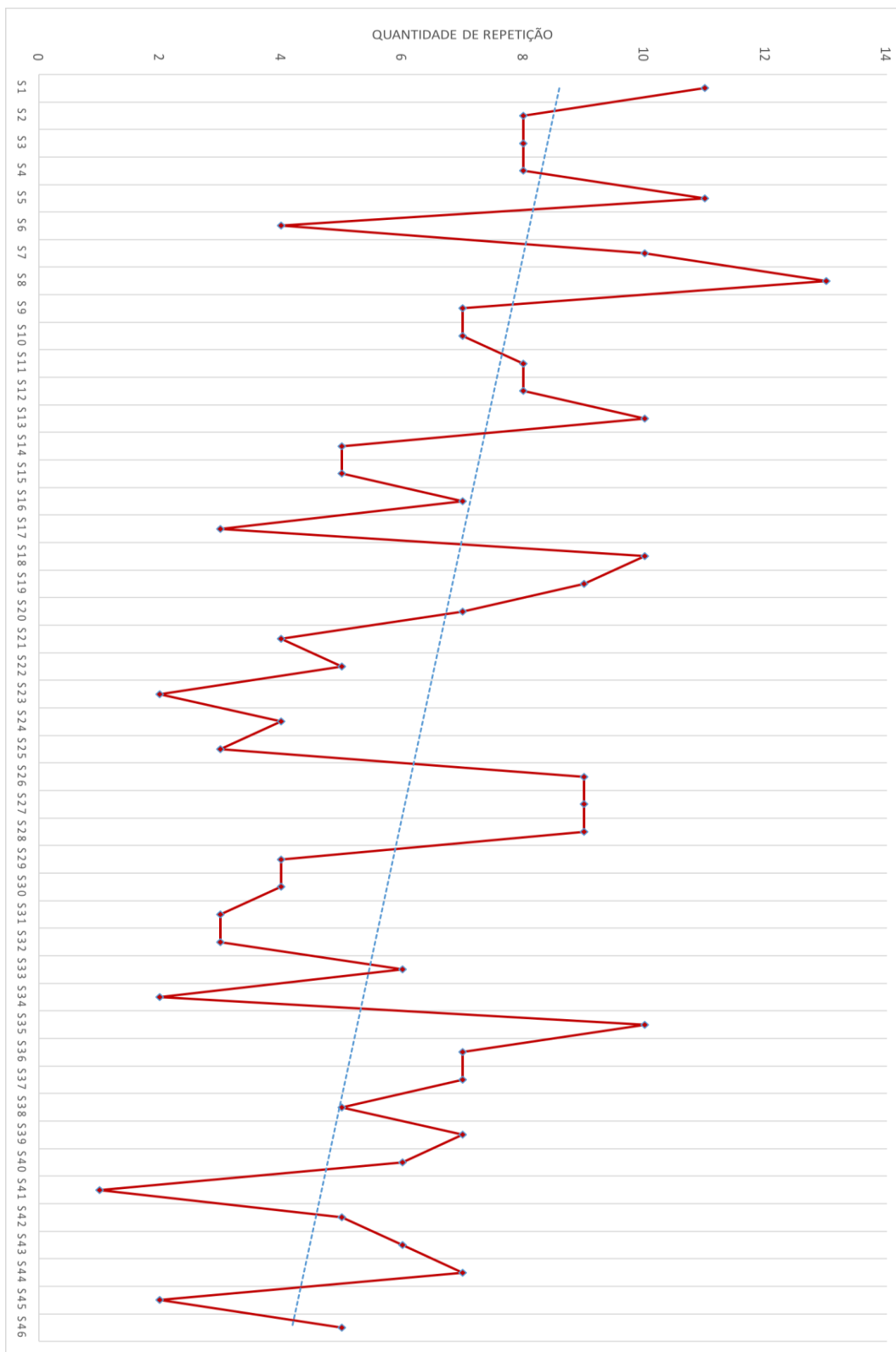
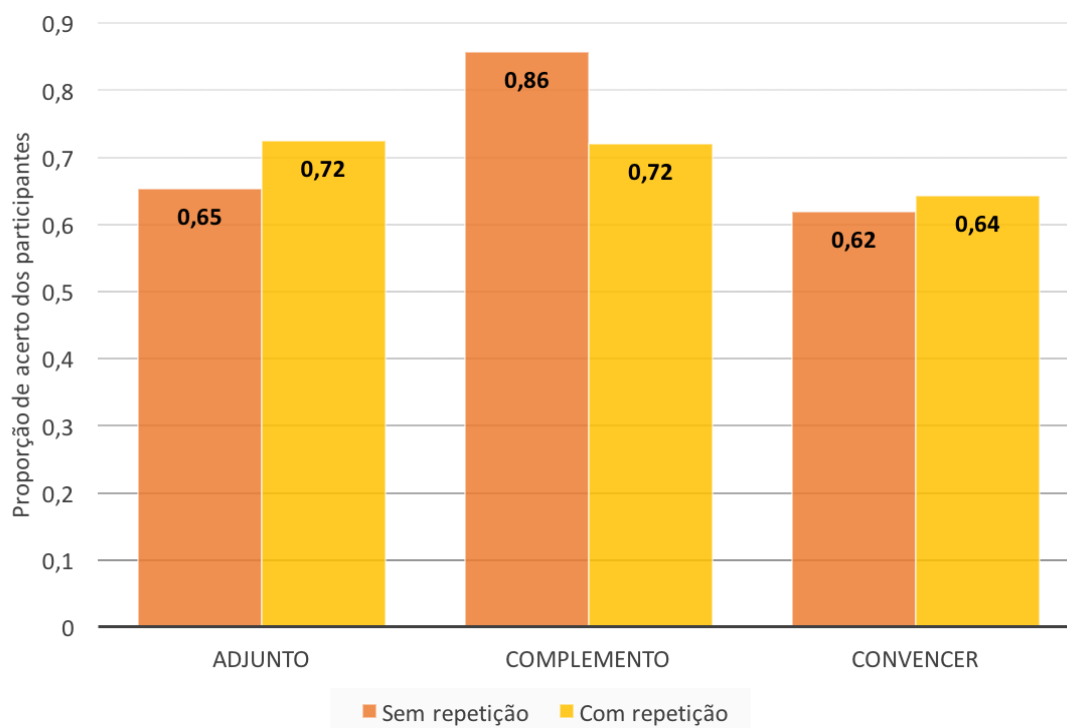


Gráfico B.26 – Proporção de acerto dos participantes nas sentenças do tipo adjunto, complemento e convencer segundo o efeito de repetir



Apêndice C – Tabelas da Análise Inferencial

Tabela C.1 - P-valor dos Testes de independência

	RESPOSTAS DO TIPO ADJUNTO	REPOSTAS DO TIPO COMPLEMENTO	RESPOSTAS DO TIPO CONVENCER	QUANTIDADE DE REPETIÇÃO
TIPO DE SINTAGMA	0,01801	0,5706	$6,801 \times 10^{-5}$	0,352

Tabela C.2 - P-valor dos Testes de correlação

		RESPOSTAS "SIM" NO ADJUNTO	RESPOSTAS "SIM" NO COMPLEMENTO
RESPOSTAS "SIM" NO CONVENCER	Coeficiente de correlação de Pearson	0,48301	0,24496
	p-valor	0,03619	0,3121
		ACERTOU NO DISTRATORA	ACERTOU NO FAMILIARIZAÇÃO
RESPOSTAS "NÃO" NO CONVENCER	Coeficiente de correlação de Pearson	0,24074	0,11299
	p-valor	0,3208	0,6451
RESPOSTAS "SIM" NO ADJUNTO	Coeficiente de correção de Pearson		0,15827
	p-valor		0,5171

Tabela C.3 - P-valor dos Testes de tendência

	ADJUNTO	COMPLEMENTO	CONVENCER	REPETIÇÃO
TESTE DE COX-STUART	0,8551	0,3613	0,8551	0,006449
TESTE DE MANN- KENDALL	0,78592	0,7836	0,20187	0,0016477

Tabela C.4 – Ajuste do modelo completo

	ESTIMATIVA	ERRO PADRÃO	VALOR Z	PR(> Z)
INTERCEPTO	2,30536	0,58708	3,927	8,61e-05
ORDEM	-0,01681	0,01334	-1.260	0,2076
TS: CONVENCER	-1,86527	0,36276	-5,142	2,72e-07
TS: DISTRATORA	0,06623	0,39571	0,167	0,8671
TS: FAMILIARIZACAO	0,78159	0,50627	1,544	0,1226
TS: ADJUNTO	-0,40716	0,39482	-1,031	0,3024
TSI: SIMPLES	-0,04285	0,24788	-0,173	0,8627
REPETICAO	-1,75170	1,01942	-1,718	0,0857

Tabela C.5 – Ajuste do modelo final

	ESTIMATIVA	ERRO PADRÃO	VALOR Z	PR(> Z)
INTERCEPTO	1,38883	0,26986	5,147	2,65e-07
TS: CONVENCER	-1,90218	0,36194	-5,255	1,48e-07
TS: DISTRATORA	0,02619	0,39700	0,066	0,9474
TS: FAMILIARIZACAO	0,87590	0,42796	2,047	0,0407
TS: ADJUNTO	-0,71818	0,35826	-2,005	0,0450

Tabela C.6 – Ajuste do modelo sem observação 30

	ESTIMATIVA	ERRO PADRÃO	VALOR Z	PR(> Z)
INTERCEPTO	1,48939	0,23531	6,330	2,46e-10
TS: CONVENCER	-2,03312	0,30961	-6,567	5,14e-11
TS: DISTRATORA	0,02887	0,34839	0,083	0,9339
TS: FAMILIARIZACAO	-0,29442	0,34914	-0,843	0,3991
TS: ADJUNTO	-0,77804	0,30840	-2,523	0,0116

Apêndice D – Diagnóstico do Modelo

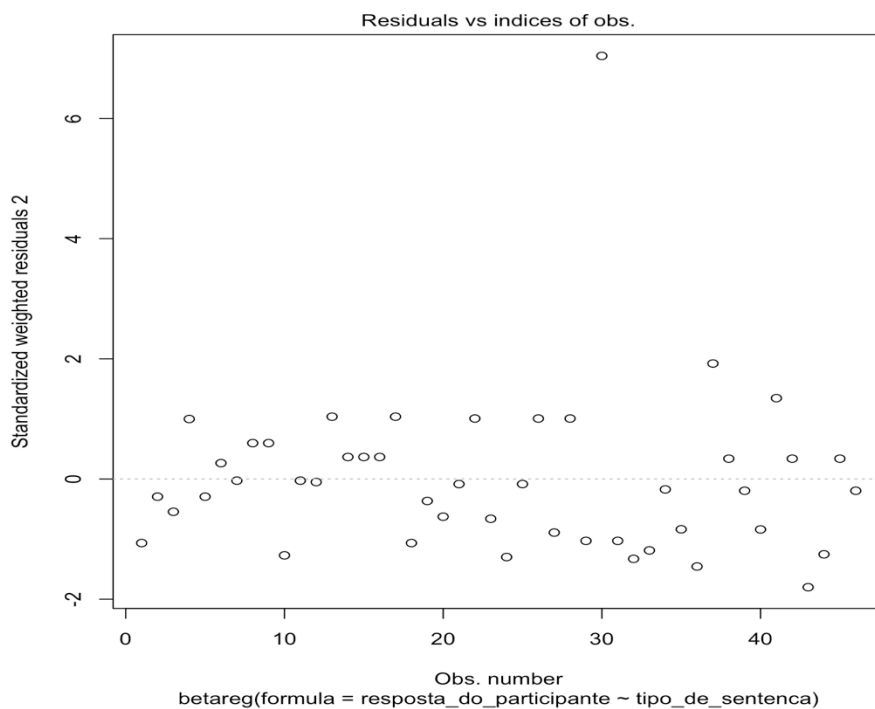
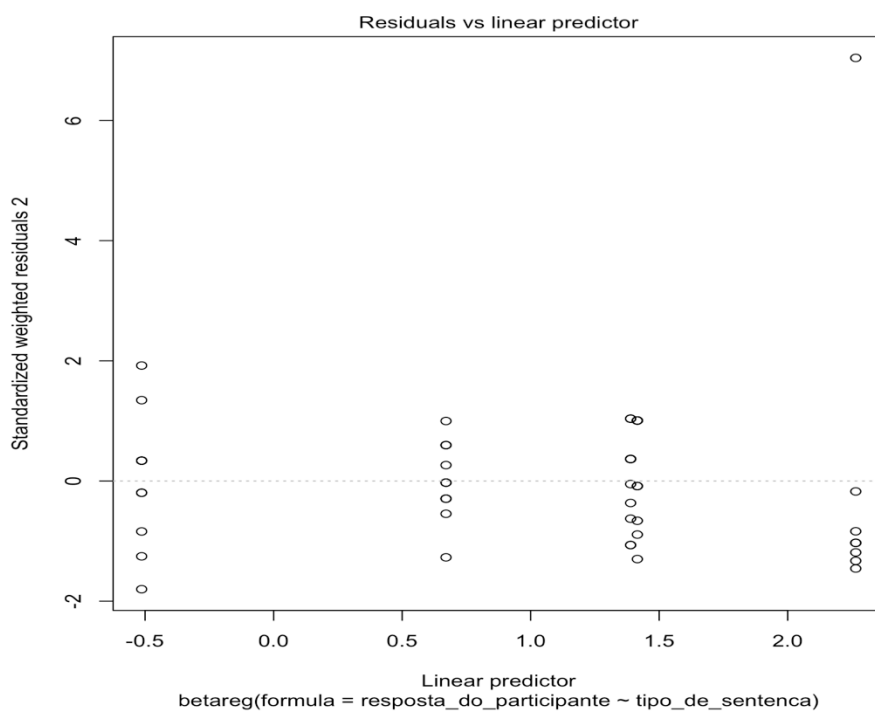
Gráfico D.1 – Resíduos x Observação para modelo final**Gráfico D.2 – Resíduos x preditor linear para modelo final**

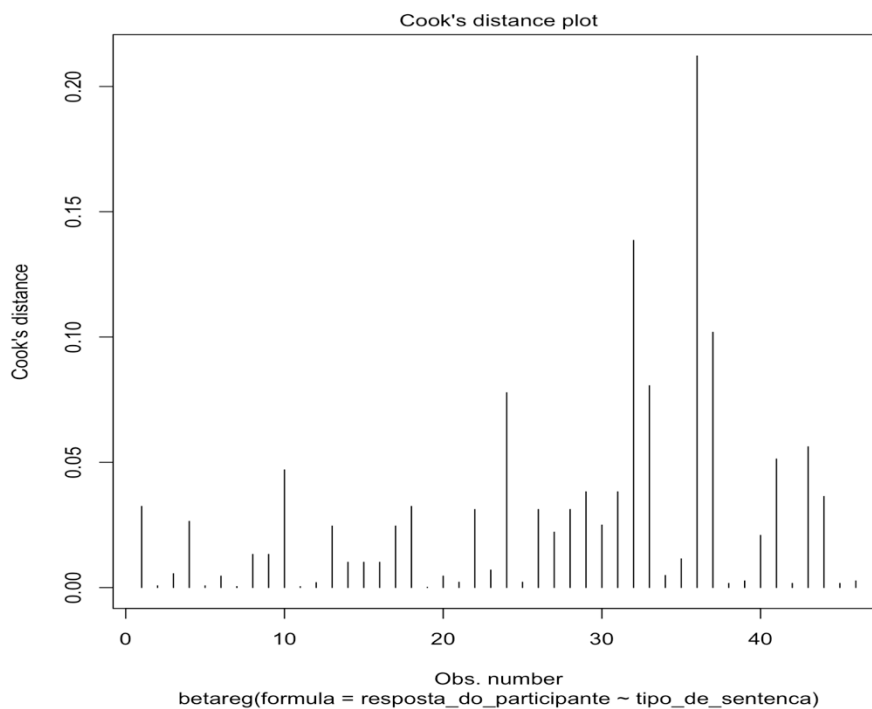
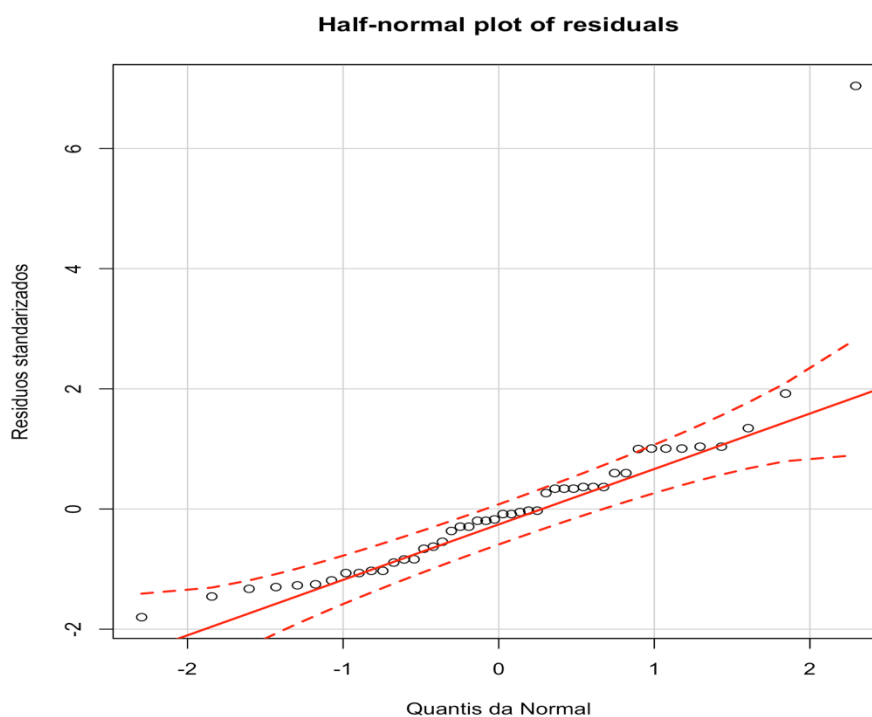
Gráfico D.3 – Distância de Cook para modelo final**Gráfico D.4 – QQ-plot para modelo final**

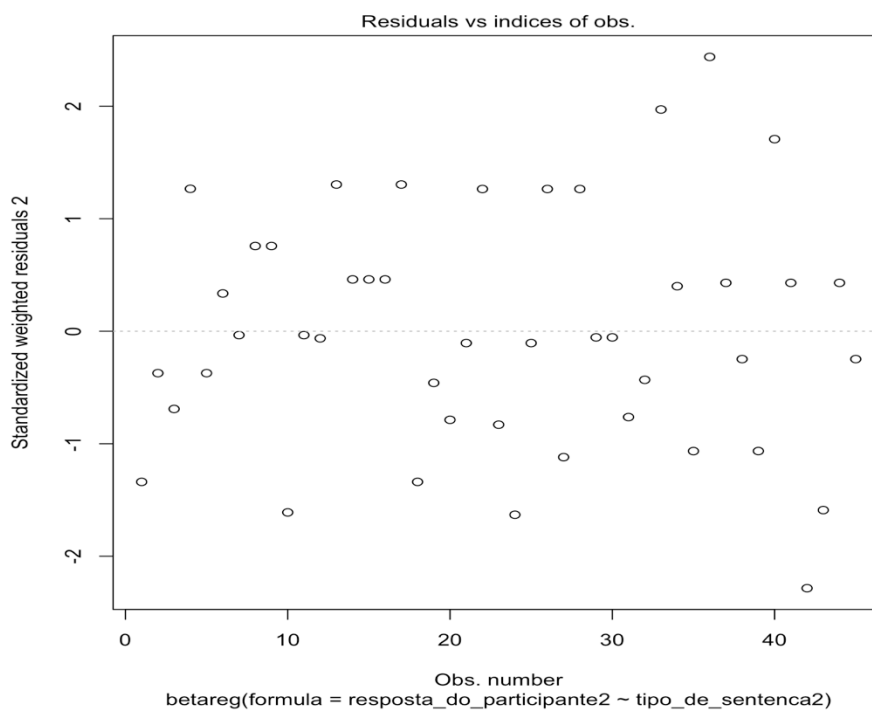
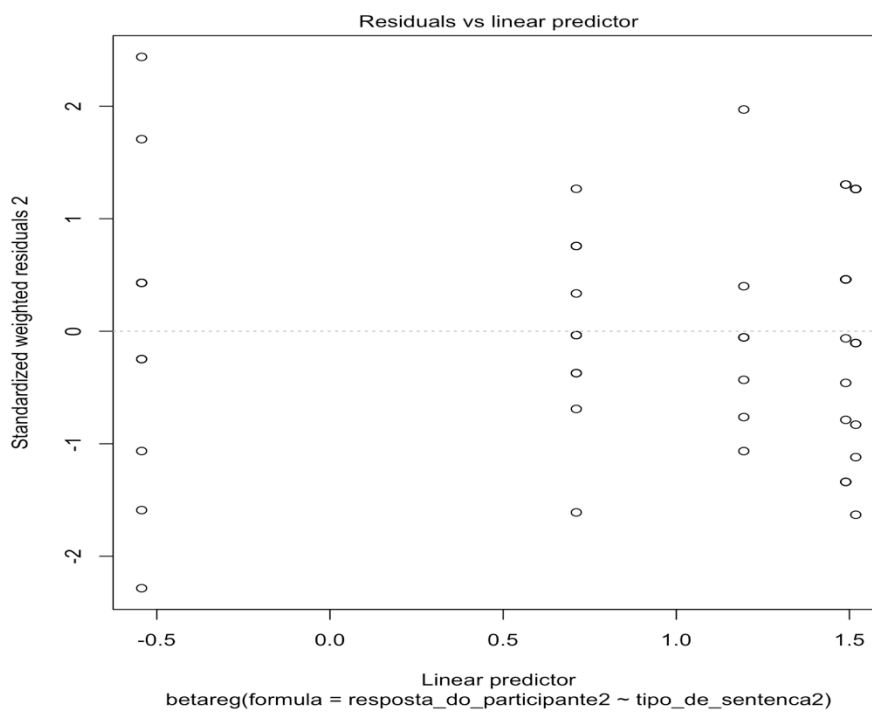
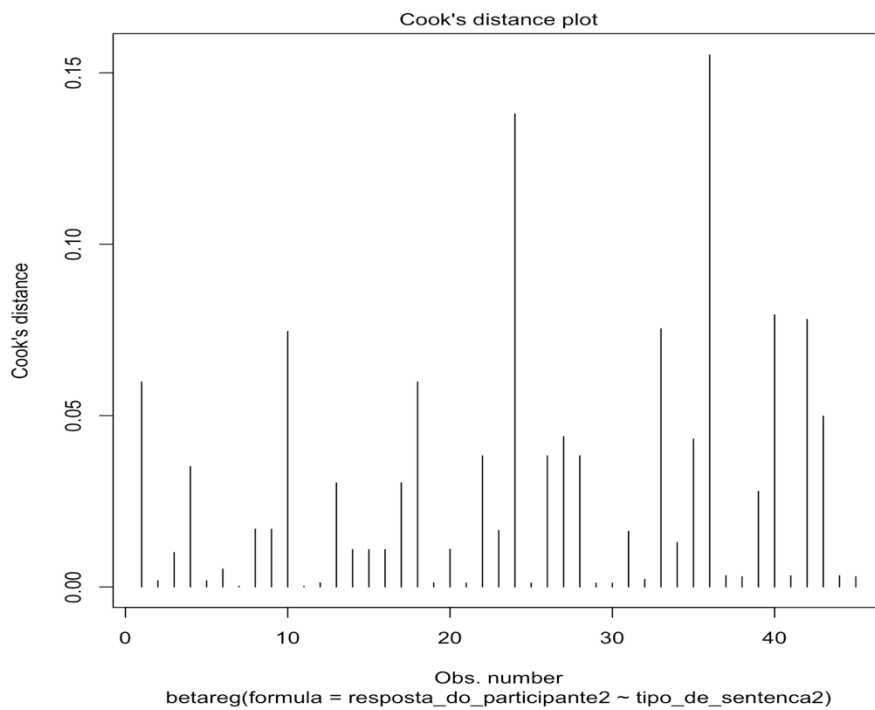
Gráfico D.5 – Resíduos x Observação para modelo sem observação 30**Gráfico D.6** – Resíduos x preditor linear para modelo sem observação 30

Gráfico D.7 – Distância de Cook para modelo sem observação 30**Gráfico D.8 – QQ-plot para modelo sem observação 30**