

MESTRADO
LINGUÍSTICA

Perceção das vogais médias anteriores pelos falantes nativos de PE

Megumi Im

M

2019



Megumi Im

**Perceção das vogais médias anteriores
pelos falantes nativos de PE**

Dissertação realizada no âmbito do Mestrado em Linguística,
orientada pelo Professor Doutor João Manuel Pires da Silva e Almeida Veloso
e coorientada pelo Professor Doutor Luís Miguel Teixeira de Jesus

Faculdade de Letras da Universidade do Porto

maio de 2019

Perceção das vogais médias anteriores pelos falantes nativos de PE

Megumi Im

Dissertação realizada no âmbito do Mestrado em Linguística,
orientada pelo Professor Doutor João Manuel Pires da Silva e Almeida Veloso
e coorientada pelo Professor Doutor Luís Miguel Teixeira de Jesus

Membros do Júri

Professora Doutora Maria de Fátima Favarrica Pimenta de Oliveira
Faculdade de Letras - Universidade do Porto

Professor Doutor Fernando Jorge da Costa de Brissos
Faculdade de Letras - Universidade de Lisboa

Classificação obtida: 18 valores

*«Do the difficult things while they are easy
and do the great things while they are small.*

A big tree begins from a tiny sprout.

A high building begins from a heap of soil.

A journey of a thousand miles begins with a single step.»

Lao-Tsé (老子)

Dedico este trabalho à minha mãe.

Sumário

Declaração de honra	9
Agradecimentos	10
Resumo	11
Abstract.....	12
Índice de ilustrações	13
Índice de tabelas	15
Lista de abreviaturas e siglas	16
0. Introdução.....	17
0.1. Formulação de questão	17
0.2. Objetos de estudo.....	20
0.3. Objetivos e hipóteses	21
0.4. Importância da investigação	21
0.5. Metodologia da investigação	22
0.6. Estrutura da dissertação	22
Capítulo 1. - A produção e a percepção da fala, e a especificidade da língua.....	24
1.1. Percepção da fala	24
1.1.1. Cadeia da fala (processo entre a produção e a percepção da fala)	24
1.1.2. Percepção da fala como categorização.....	25
1.1.3. Especificidade da língua.....	27
1.1.4. Aquisição de categorias e a alteração da capacidade perceptiva	30
1.2. Produção e percepção das vogais.....	31
1.2.1. Produção das vogais	31
1.2.2. As propriedades acústicas e a percepção das vogais.....	32
1.3. Teoria do protótipo: estrutura interna da categoria dos sons da fala	35
1.4. Teoria do protótipo na base de experiências da percepção da fala com <i>goodness rating</i>	37
Capítulo 2. – Influência do sistema fonológico e as vogais do PE.....	39

2.1. Distinção entre o nível fonológico e fonético	39
2.2. Influência do sistema fonológico na percepção da fala	40
2.3. Propostas teóricas do inventário fonológico das vogais do PE	42
2.4. Aquisição do processo fonológico (neutralização) do PE pelas crianças	48
2.5. Variações/ alternâncias no nível fonético	50
2.6. Características fonético-acústicas das vogais médias anteriores de Lisboa.....	52
Capítulo 3. - Os estudos experimentais sobre a influência das relações fonológicas na percepção da fala	54
3.1. As relações fonológicas e a percepção da fala.....	54
3.2. Criação dos estímulos nas experiências da percepção de fala	61
3.3. <i>TANDEM-STRAIGHT</i>	62
Capítulo 4. - Metodologia.....	65
4.1. Estímulos	65
4.1.1. Seleção da voz de Lisboa	65
4.1.2. Inquérito online	66
4.1.3. Gravação das vogais-alvo.....	69
4.1.4. Criação dos estímulos	69
4.2. Questionário.....	72
4.3. Autorização	72
4.4. Pré-teste	73
4.4.1. Participantes do pré-teste	73
4.5. Teste Principal	73
4.5.1. Participantes	73
4.6. Procedimento e <i>design</i> da experiência.....	73
4.7. Tratamento dos dados	75
Capítulo 5. - Resultados	76
5.1. Tarefa de identificação.....	76
5.2. Fronteiras entre categorias	80

5.3. Tarefa de classificação de <i>goodness</i>	82
5.3.1. Estrutura interna de categorias e protótipos	84
5.3.2. Comparação de grau de nitidez entre as fronteiras	87
Capítulo 6. – Discussão dos resultados	93
6.1. As quatro categorias vocálicas.....	93
6.2. Estrutura interna da categoria vocálica.....	94
6.3. Grau de nitidez das fronteiras	95
6.4. Influência das relações fonológicas na percepção da fala: Suposição de Troubetzkoy (1969 [1939]).....	99
6.5. <i>Perceptual hyperspace</i> (Johnson, 2000) e <i>adaptive dispersion</i> (Liljencrants & Lindblom, 1972)	100
Capítulo 7. – Conclusão e considerações finais	101
7.1. Conclusões gerais	101
7.2. Limitações da investigação	102
7.3. Trabalhos Futuros	104
Referências bibliográficas	106
Anexos.....	117
Anexo 1 – Questionário sobre o perfil biográfico e linguístico dos participantes....	117
Anexo 2– Autorização	118
Anexo 3– Dados biográficos, linguísticos, sociodemográficos e comentários dos participantes no pré-teste	120
Anexo 4– Dados biográficos, linguísticos, sociodemográficos e comentários dos participantes na experiência principal.....	121
Anexo 5– Gráficos: Teste principal	123
Anexo6– Gráficos: Pré-teste	143

Declaração de honra

Declaro que a presente dissertação é de minha autoria e não foi utilizado previamente noutro curso ou unidade curricular, desta ou de outra instituição. As referências a outros autores (afirmações, ideias, pensamentos) respeitam escrupulosamente as regras da atribuição, e encontram-se devidamente indicadas no texto e nas referências bibliográficas, de acordo com as normas de referenciação. Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.

Braga, abril de 2019

Megumi Im

Agradecimentos

Ao Professor Doutor João Veloso, meu orientador, pelos comentários e sugestões pertinentes para a realização do trabalho.

Ao Professor Doutor Luís Jesus, meu coorientador, por ter aceitado coorientar este trabalho e pelos seus apoios fundamentais ao mesmo.

À Professora Doutora Maria da Conceição Costa da Universidade de Aveiro pela palavra amiga de incentivo e pela sua disponibilidade para resolver questões estatísticas.

À Professora Doutora Maria Aldina Marques e ao Professor Doutor José Teixeira da Universidade do Minho pela sua inspiração para a *Introdução* deste trabalho.

Aos participantes que realizaram as tarefas na recolha de dados e àqueles que participaram na gravação para criar estímulos.

Ao André Amorim, André Trindade, Catarina Capela e sua mãe pelo fornecimento dos contatos lisboetas.

Aos peritos da fonética e/ou da dialetologia do português europeu que participaram no inquérito online para a escolha de um falante.

Ao Rodrigo Pereira pela correção do meu português ao longo da escrita desta dissertação. E também à Beatriz Fachada pela ajuda, na fase final do trabalho, com as últimas revisões

E por último, à minha mãe e aos meus amigos, que me apoiaram em todos os momentos, pelo estímulo e compreensão.

A todos os que contribuíram de algum modo para a elaboração deste trabalho, o meu mais sincero agradecimento.

Resumo

O reflexo do sistema fonológico da primeira língua (L1) na percepção da fala da segunda língua (L2) tem sido explorado na área da aquisição da língua segunda. Esta influência, no entanto, pode também ser observada na percepção dos sons da L1, de acordo com as relações que os fonemas do seu sistema estabelecem (Troubetzkoy, 1939 [1969]).

Assim, este estudo tem como objetivo averiguar a influência da relação de neutralização na percepção dos falantes nativos de português europeu (PE), analisando o grau de nitidez das fronteiras entre categorias fonológicas. As vogais médias anteriores foram escolhidas como vogais-alvo, devido à complexidade das suas relações no sistema do PE (Veloso, 2016). Nesta investigação, optou-se pela visão da teoria do protótipo (Roch, 1975), postulando que a categorização é o mecanismo cognitivo fundamental na percepção dos sons da fala (Holt & Lott, 2010).

Quinze falantes (sexo feminino = 9) nativos de PE, oriundos de Lisboa, licenciados ou a frequentar uma licenciatura, com idades entre 19 e 34 anos (média = 23.7 anos) participaram numa experiência constituída por duas tarefas: identificação e classificação de *goodness*, com 31 estímulos do continuum /i/-/a/.

Através da análise estatística dos dados obtidos, confirmou-se uma menor nitidez na fronteira entre /e/ e /ɛ/ do que entre /ɛ/ e /a/, sendo o último contraste não neutralizável. Assim, confirma-se a existência de uma possível influência das relações fonológicas na percepção dos falantes de PE. Contudo, o menor grau de nitidez não foi atribuído, estatisticamente, à fronteira entre /e/ e /ɛ/, mas entre /i/ e /e/, o que não coincide com um estudo anterior acerca do português do Brasil (PB) (Silva & Neves, 2016). Este resultado poderá ser devido a diferentes tipos de neutralização entre PE e PB, à especificidade dialetal ou a diferentes metodologias entre as duas investigações.

Palavras-chave: percepção da fala, influência do sistema fonológico, vogais médias anteriores, categorização, Teoria do protótipo

Abstract

The reflection of the first language's (L1) phonological system on the second language's (L2) perception has been explored on the field of second language acquisition. This influence, however, can also be observed on L1 sounds' perception, according to how the system's phonemes establish relations (Troubetzkoy, 1939 [1969]).

As such, this study's purpose is to ascertain the influence of the neutralization relation on the perception of European Portuguese (PE) speakers, analyzing the degree of boundary sharpness between phonological categories. The mid front vowels were chosen as target vowels due to their relationships' complexity on PE's system (Velooso, 2016). In this investigation, we worked under the light of Prototype Theory (Roch, 1975), assuming that categorization is the fundamental cognitive mechanism in the perception of speech sounds (Holt & Lott, 2010).

Fifteen speakers (female = 9), PE natives from Lisbon, either already graduated from university or currently attending it, between the ages of 19 and 34 (average = 23.7) participated in an experiment composed of two tasks: identification and goodness classification, with 31 stimuli from the /i/-/a/ continuum.

Through statistical analysis of the obtained data, less sharpness was confirmed in the boundary between /e/ and /ɛ/, when comparing with the one between /ɛ/ and /a/, with the latter being a non-neutralizable contrast. As such, a possible influence of phonological relations on PE speakers' perception was concluded. Nevertheless, this lesser degree of sharpness was not statistically attributed to the boundary between /e/ e /ɛ/ but to the one between /i/ e /e/, which does not agree with a previous study about Brazilian Portuguese (PB) (Silva & Neves, 2016). This result could be due to different types of neutralization between PE and PB, to dialectal specificity or to the different methodologies employed in both investigations.

Keywords: speech perception, influence of phonological system, mid front vowels, categorization, the prototype theory

Índice de ilustrações

Figura 1. Processos envolvidos na produção (da esquerda para a direita) e compreensão (da direita para a esquerda) da linguagem humana. Adaptado de Fernández e Carins (2011: 16).....	24
Figura 2: Forma de onda para a frase “No meu caso deve ser uma exceção”	26
Figura 3: Espectrogramas de banda larga das nove vogais orais do português europeu padrão, produzidas por uma falante feminina. Adaptados de Falé (2005:125)	33
Figura 4: Representações gráficas dos dados das tarefas de identificação e de discriminação das categorias das vogais [i]-[ɪ] e das consoantes [b]-[p] (Pisoni, 1973). Adaptado de reprodução por Cutler (2012: 9).....	35
Figura 5: O espaço vocálico de 4 grupos: PB masculino e feminino; PE masculino e feminino. Adaptado de Escudero et al. (2009: 1385).....	53
Figura 6: Os quatro tons do mandarim em espaços perceptivos para os ouvintes de chinês mandarim e de inglês americano. Adaptado de Hume & Johnson (2003).....	55
Figura 7: Espectros da vogal /ɛ/ gerados pelo STRAIGHT (Kawahara et al. 1999) e pelo sintetizador Klatt (1980). Adaptado de Liu & Kewley-Port (2004: 1231).	63
Figura 8: Captura de ecrã das páginas de avaliação do inquérito online.	67
Figura 9: Formas de onda (em cima) e espectrogramas (em baixo) das quatro amostras das vogais orais do português europeu (de Lisboa), produzidas por um falante masculino	70
Figura 10: Captura de ecrã da tarefa de identificação e de classificação de goodness	74
Figura 11: As curvas de regressão logística de todos os participantes para a transição da categoria /i/ para não /i/.....	78
Figura 12: As curvas de regressão logística de todos os participantes para a transição da categoria /e/ para não /e/.	78
Figura 13: As curvas de regressão logística de todos os participantes para a transição da categoria /ɛ/ para não /ɛ/.	79
Figura 14: A distribuição dos valores das três fronteiras, /i/-/e/, /e/-/ɛ/ e /ɛ/-/a/ para os 15 participantes.	80
Figura 15: Representação do goodness rating mediano de todos os participantes.	83
Figura 16: Representação do goodness rating do informante CC.....	84
Figura 17: Representação do goodness rating do informante LP	85
Figura 18: Representação do goodness rating do informante BM.....	86
Figura 19: Representação do goodness rating do informante AC	86
Figura 20: Representação do goodness rating do informante CL.....	87
Figura 21: A frequência absoluta de respostas X para cada um dos 31 estímulos	88

Figura 22: A frequência absoluta de respostas X no pré-teste para cada um dos 31 estímulos	96
Figura 23: Representação do goodness rating mediano dos 6 informantes do Norte de Portugal.	98

Índice de tabelas

Tabela 1: Classificação das vogais orais do Português Europeu padrão. Adaptado de Falé (2005: 79).....	32
Tabela 2: As realizações fonéticas de /ɛ/ no contexto antes de consoante palatal, pelos falantes do dialeto de Lisboa. Adaptado de Rodrigues (2003: 129).	51
Tabela 3: As realizações fonéticas de /e/ no contexto antes de consoante palatal, pelos falantes do dialeto de Lisboa. Adaptado de Rodrigues (2003: 130).	51
Tabela 4: Dados sociodemográficos dos falantes.	66
Tabela 5: Avaliação das gravações pelos 8 peritos da fonética e da dialetologia do PE. .	68
Tabela 6: Valores médios de f ₀ , F ₁ , F ₂ e F ₃ (Hz) e duração (ms) das vogais [i],[e],[ɛ] e [a] produzidas pelos falantes masculinos de Lisboa no estudo de Escudero <i>et al.</i> (2009: 1383)	70
Tabela 7: Valores de f ₀ , F ₁ , F ₂ e F ₃ (Hz) e duração (ms) dos 31 estímulos sintetizados. .	71
Tabela 8: Fronteiras médias estimadas entre categorias e os valores dos formantes correspondentes.....	79
Tabela 9: Valores de <i>p</i> dos testes de normalidade para as diferentes amostras	81
Tabela 10: Resultado das comparações múltiplas entre grupos (perspetiva não paramétrica e paramétrica).....	82
Tabela 11: Os valores individuais e os medianos das fronteiras. Os medianos são: /i/-/e/=5.65, /e/-/ɛ/=14.46, /ɛ/-/a/ =24.95.	89
Tabela 12: Valores de <i>p</i> dos testes de normalidade para as diferentes amostras	90
Tabela 13: Testes de comparação de medianas/ médias para amostras emparelhadas.....	91
Tabela 14: Resultado do cálculo do rácio entre goodness scores nas três fronteiras.....	92

Lista de abreviaturas e siglas

2AFC – Two alternatives forced choice

4I2FC – Four intervals two alternatives forced choice

AD – Teste de Anderson-Darling

ANOVA – Análise de variância

KS – Tste de Kolmogorov-Smirnov

L1 – Língua Materna

L2 – Língua Segunda

PB – Português do Brasil

PE – Português Europeu

QECR – Quadro Europeu Comum de Referência para as Línguas

UPSID – The UCLA Phonological Segment Inventory Database

VOT – Voice Onset Time

0. Introdução

0.1. Formulação de questão

Enquanto seres humanos, tendemos a pensar que temos, pelo menos, uma certa ideia da vastidão imensamente diversa de culturas que permeiam o nosso planeta. No entanto, não conseguimos deixar de nos surpreender com coisas muitas vezes tão básicas como as diferenças na linguagem. Por exemplo, certamente estranharíamos o facto de os esquimós usarem sete palavras diferentes para se referirem a diferentes tipos de neve, ou ainda que, na Ilha de São Lourenço, são utilizadas diariamente 99 palavras para indicar espécies de gelo distintas (Miyaoka, 2006: 45-46). No sentido oposto, temos também línguas como o vietnamita, onde não existem duas palavras distintas para falar das cores azul e verde (Alvarado, & Jameson, 2002 :71). Tudo isto nos faz questionar o quão diferente a nossa visão do mundo é da daqueles que vivem a milhares de quilómetros de nós e comunicam usando línguas que para nós são completamente incompreensíveis. Além destas questões semânticas, é também notável a variabilidade dos sons da fala. Notamos isto particularmente quando aprendemos uma língua estrangeira numa fase mais tardia da vida, pois encontramos dificuldades e até mesmo a incapacidade de distinguir certos sons que lhe pertencem. Aliás, é bastante conhecido, na literatura, o exemplo dos japoneses, que não são capazes de distinguir entre as consoantes [r] e [l], que aos portugueses soam perfeitamente diferentes (Miyawaki *et al.*, 1975; MacKain, Best & Strange, 1981; Mochizuki, 1981). Então, sendo nós parte da mesma espécie, possuindo todos, em geral, os mesmos órgãos sensoriais, impõe-se a questão: qual a razão destas diferenças na perceção? A resposta dada pela etnolinguística, pela psicolinguística e pela linguística cognitiva para as duas primeiras questões é a seguinte: numa comunidade linguística, são seleccionados os conceitos que são culturalmente mais relevantes (Miyaoka, 2006: 46), sendo denominados com palavras (Miyaoka, 2006: 44-45). Cada língua, então, cria a sua própria organização desses conceitos, e constrói o mundo mental dos seus falantes (Silva, 1997: 61). Deste modo, a língua/linguagem interfere cognitivamente connosco, pois o nosso conhecimento do mundo é desenvolvido desde a infância, conhecimento esse que é capturado através da nossa língua nativa (Miyaoka, 2006: 49). No entanto, a ausência de certas palavras no léxico de uma língua não significa que, por exemplo, os portugueses não consigam identificar diversos tipos de neve

ou que os vietnamitas não percebiam a diferença entre o azul e o verde¹ (Brown, 1954: 458). O que está em causa é a presença no seu léxico de palavras que dividem o mundo em categorias conceituais distintas.

No que toca à terceira questão, a sua resposta não está relacionada com a cultura, mas sim com o sistema linguístico. Ou seja, essa dificuldade dos japoneses tem origem no sistema fonológico, uma vez que estas consoantes pertencem à mesma categoria no sistema da língua japonesa (Miyawaki *et al.*, 1975; MacKain, Best & Strange, 1981; Mochizuki, 1981). Assim, embora os domínios linguísticos dos exemplos sejam diferentes, vemos que o que têm em comum é a categorização.

Os sons relevantes no sistema linguístico foram denominados por Troubetzkoy (1969 [1939]) como fonemas. No seu modelo teórico da fonologia, os fonemas possuem uma função distintiva numa dada língua e são entidades abstratas, armazenadas no conhecimento da língua dos falantes. Estes são ativados na produção e perceção dos sons da fala (Veloso, 2015: 8). Durante a aquisição da primeira língua, as crianças aprendem os fonemas do seu sistema linguístico, agrupando os fones (sons da fala) altamente variáveis em categorias. Além disso, aprendem também regras fonológicas que referem quais os diferentes sons que têm os mesmos valores no nível subjacente (Kuhl, 2004: 832). Neste processo da aquisição, as crianças, que conseguiam, inicialmente, distinguir os sons de qualquer fala (Werker & Tees, 1984a; Werker & Lalonde 1988; Kuhl, 2004) “desaprendem” os sons que não pertencem ao sistema da sua L1, (Werker & Tees, 1984a, Werker & Lalonde, 1988; Tsushima *et al.* 1994; Kuhl, 2004; Kuhl *et al.*, 2005) para assim melhor assimilarem a sua língua materna (Kuhl *et al.*, 2005). Deste modo, a nossa perceção da fala fica formatada conforme o sistema da primeira língua (Werker & Tees, 1984a; Kuhl *et al.*, 2008), demonstrando, assim, uma influência maior na perceção, relativamente ao domínio da semântica.

Por outro lado, a influência do sistema fonológico na perceção pode ser observada mesmo nos sons existentes do sistema da língua nativa (Troubetzkoy 1969 [1939]; Hume & Johnson, 2003; Boomershine *et al.* 2008). No que diz respeito à possível influência das relações fonológicas na perceção dos sons da língua nativa, demonstraremos em seguida a ideia de Troubetzkoy, em *Grundzüge der Phonologie* (1969 [1939]),

¹ Em vietnamita existem as expressões para referir azul e verde: **xanh** biêng (xanh “da pele do céu”) e **xanh** lá cây (xanh “da folha de árvore”) (cf. Phụng, 1998: 59 e 236). Assim sendo, podemos dizer que estas são as subcategorias da categoria de uma cor “xanh”. (cf. Alvarado & Jameson, 2002: 71).

“The psychological difference between constant and neutralizable distinctive oppositions is very great. Constant distinctive oppositions are perceived clearly even by those members of the speech community who have had no phonetic training. The terms of such an opposition are considered two distinct “phonetic entities”. In neutralizable distinctive oppositions perception fluctuates: in positions of relevance both opposition members are clearly distinguished; in positions of neutralization, on the other hand, it is often not possible to indicate which of the two had just been produced or perceived. However, even in the position of relevance, members of a neutralizable opposition are often felt only as two-meaning-differentiating nuances, that is, as two distinct yet closely related phonetic entities.” (Troubetzkoy, 1969 [1939]: 78)

Assim sendo, surgiu a seguinte questão: de que forma é que os nativos portugueses percebem as vogais médias anteriores que estabelecem uma relação de neutralização no seu sistema linguístico?

Considerando a categorização como o processo cognitivo subjacente e fundamental na percepção da fala (Holt & Lott, 2010), deve-se, primeiramente, escolher uma teoria que a aborde. Recuando no tempo, o modelo clássico da categorização já se encontrava no pensamento aristotélico, sendo posteriormente retomado e desenvolvido pela semântica estruturalista (Ferrari, 2011: 33). Segundo este modelo, cada elemento pertencente a uma dada categoria deverá possuir todos os traços que definem essa categoria e, por conseguinte, os membros de uma categoria são homogêneos e as categorias demonstram limites rígidos (Silva, 1997: 63; Ferrari, 2011: 33). O modelo clássico, contudo, não conseguiu resolver algumas questões e motivou, nos anos 70, os investigadores da área da psicologia e da sociolinguística a realizarem estudos experimentais com o objetivo de averiguar o modo como os seres humanos categorizam as cores (Berlin & Kay, 1969; Rosch, 1973) e os objetos (Labov, 1973; Rosch, 1973). Os resultados revelaram que os membros de uma categoria perceptual não são homogêneos, nem demonstram limites rígidos entre categorias. Pelo contrário, observou-se uma construção da categoria, por parte dos membros, de uma forma gradual, ao redor do membro representativo, chamado o protótipo dessa categoria (Rosch, 1973), e também se observou que os limites eram difusos (Labov, 1973). A partir destes resultados, Rosch (1975) formulou uma proposta relativa à categorização, “a teoria do protótipo”. De acordo com esta teoria, o protótipo pode servir como referência cognitiva na decisão da categorização dos itens julgados (Rosch, 1975). Nas décadas seguintes, investigadores incentivados por esta teoria realizaram estudos experimentais

na área da fonética (Oden & Massaro, 1978; Samuel, 1982; Grieser & Kuhl, 1989; Miller & Volaitis, 1989; Kuhl, 1991; Kuhl *et al.*, 1992; Lively & Pisoni, 1997). Confirmou-se, então, a estrutura interna das categorias fonéticas ao redor do som protótipo (tal como verificado nas categorias das cores e dos objetos). Observaram-se, ainda, protótipos distintos da mesma categoria fonética, conforme a primeira língua dos informantes (Eerola *et al.*, 2014; Masapollo *et al.*, 2017).

0.2. Objetos de estudo

Partindo da pressuposição de Troubetzkoy (1969: 87[1939]), isto é, que nem todas as relações das oposições no sistema têm o mesmo valor, podendo refletir-se na percepção conforme o tipo da relação, o presente estudo visa averiguar este fenómeno na percepção dos nativos do PE. As vogais médias anteriores são escolhidas como segmentos-alvo devido à sua complexidade verificada e/ou proposta nos estudos que iremos rever no Capítulo 2: *Fonética e fonologia das vogais do PE*. Nomeadamente:

- 1) os inventários fonético e fonológico, com as suas relações multívocas, propostos por Veloso (2016);
- 2) o facto de que o tipo de neutralização das vogais médias anteriores do PE desenvolve outro termo [ɨ] que, simultaneamente, possui estatuto teórico fonológico (Veloso, 2010, 2016);
- 3) a aquisição tardia desta relação do sistema (Freitas, 2004), que corrobora a hipótese da sua complexidade;
- 4) as variações fonéticas do dialeto de Lisboa relativas ao /e/ (Barros, 1994; Rodrigues, 2003) que nos levam a pensar na possibilidade desta forma teórica ser menos consistente no sistema (pelo menos neste dialeto), sendo o sistema construído pelas relações de oposição;
- 5) ainda no mesmo dialeto a convergência de [e] e [ɛ] observada no espaço das vogais poderá ser um índice da fusão no processo da mudança (Escudero *et al.*, 2009).

0.3. Objetivos e hipóteses

O presente estudo baseia-se em dois pressupostos teóricos: a percepção dos sons da fala é influenciada pelas relações que estes sons estabelecem no seu sistema da língua (Troubetzkoy (1969 [1939]; Hume & Johnson, 2003; Boomershine *et al.* 2008); as categorias dos sons da fala, sobretudo das vogais, possuem uma estrutura interna (Kuhl, 1991; Kuhl *et al.*, 1992). Posto isto, este estudo tem como objetivo geral confirmar a influência das relações fonológicas de L1 na percepção dos falantes nativos com a metodologia que engloba o ponto de vista da teoria do protótipo. Assim sendo, delineou-se como objetivo principal explorar se a relação de neutralização estabelecida por fonemas vocálicos tem influência na percepção dos nativos do PE, mais especificamente, averiguar o grau de nitidez da fronteira entre as categorias das vogais médias anteriores definida pelos ouvintes portugueses, comparando-o com o grau das fronteiras entre outras categorias vocálicas não recuadas. Postulando uma estrutura interna da categoria fonética e, por consequente, uma difusão das fronteiras entre categorias, o grau de nitidez das fronteiras na percepção dos nativos verificar-se-á com o grau do conflito na classificação da qualidade do som. Assim sendo, de modo a atingir o nosso objetivo analisaremos: a existência de quatro categorias vocálicas não recuadas e as suas fronteiras; a presença, nestas categorias, de uma estrutura interna com o(s) seu(s) protótipo(s) e finalmente, o grau de nitidez das fronteiras.

Formulamos, então, as seguintes hipóteses:

1. Existência de quatro categorias fonémicas com uma estrutura gradual (embora possa haver dispersão intra e inter-ouvintes tal como confirmado na produção);
2. O grau de nitidez da fronteira entre as vogais médias /e/ e /ɛ/ é menor do que outras fronteiras.

0.4. Importância da investigação

No que concerne aos estudos experimentais da percepção da fala do PE, encontram-se alguns na área da aquisição da segunda língua com aprendentes estrangeiros (Shu, 2014; Oliveira, 2016). Relativamente à percepção dos falantes nativos desta língua, esta não é suficientemente explorada. Entre esses estudos, por exemplo, Castro (1993) confirmou na

percepção dos nativos do PE o paradigma categorial num continuum do vozeamento das consoantes, empregando as tarefas da identificação e da discriminação com estímulos gerados a partir da fala natural. É inédita, portanto, uma investigação com a metodologia que engloba o ponto de vista da teoria do protótipo que visa verificar a influência das relações do sistema fonológico com a percepção dos falantes nativos de PE.

Deste modo, a realização da presente investigação poderá contribuir com dados empíricos para a compreensão global, não apenas do sistema desta língua, mas também das línguas em geral. É pertinente salientar que estamos conscientes de que os resultados deste trabalho não responderão diretamente às questões dos modelos fonológicos. Esta investigação será apenas uma tentativa de produzir evidência empírica dos comportamentos cognitivos dos nativos do PE, mais concretamente, do dialeto de Lisboa², relativos às vogais em questão.

0.5. Metodologia da investigação

O método de investigação escolhido foi o estudo experimental, recorrendo-se à experiência como instrumento de recolha de dados. Estes dados foram analisados estatisticamente para caracterizar o processo da categorização dos nativos do PE no que respeita às vogais em questão. As experiências decorreram entre março e junho de 2018 em Lisboa, envolvendo jovens lisboetas que satisfizeram os critérios de inclusão.

0.6. Estrutura da dissertação

A presente dissertação está estruturada em 7 capítulos, antecedidos de uma introdução.

O capítulo 1, *A percepção da fala e a especificidade da língua*, aborda, de forma geral, os mecanismos na produção e na percepção da fala, de modo a contextualizar o nosso estudo na cadeia da fala. Seguidamente, apresenta-se o processo cognitivo fundamental e universal na

² Temos consciência de que não é possível extrapolar quaisquer resultados ou descrições para o português em todas as suas variedades, nem tão pouco para uma só variedade, através das amostras desta investigação, cujas dimensão e características possuem certas particularidades (com destaque para a formação universitária e a faixa etária de todos os seus informantes). Embora, do ponto de vista variacionista, o português falado pelos sujeitos da investigação não corresponda a uma amostra homogénea, iremos referir ao longo de estudo as suas características sociodialetais como "variedade da amostra".

percepção da fala, a categorização, que nos permite obter informação acerca da especificidade da língua. É também referido o processo da aquisição dos sons da primeira língua, que vai alterar a percepção da fala. O capítulo termina com as descrições das propriedades acústicas da produção e da percepção das vogais de forma a introduzir a teoria norteadora da nossa investigação.

O capítulo 2, *Influência do sistema fonológico e as vogais do PE*, apresenta o conceito básico da distinção entre o nível fonológico e fonético, e uma linha de investigação que visa incorporar a percepção da fala na compreensão global do sistema dos sons da fala. Abordam-se, de seguida, algumas propostas a respeito da descrição fonológica das vogais do PE, referindo o problema do estatuto fonológico das vogais-alvo neste trabalho, que pode ter impacto na percepção dos nativos. São também relatados alguns estudos relacionados com estas vogais de maneira a sustentar a nossa questão.

O capítulo 3, *Os estudos experimentais sobre a influência das relações fonológicas na percepção da fala*, apresenta, com uma visão crítica, os mais recentes estudos relevantes para a nossa investigação.

O capítulo 4, *Metodologia*, apresenta o processo de investigação, descrevendo o método usado, referindo a técnica de recolha de dados, a definição e a caracterização da amostra da população e os critérios adotados de seleção da população.

No capítulo 5, *Resultados*, apresentam-se e analisam-se os resultados dos dados recolhidos durante a investigação.

No capítulo 6, *Discussão dos resultados*, interpretam-se e discutem-se os resultados apresentados no capítulo anterior, de modo a dar respostas à questão colocada e validar as hipóteses formuladas, para assim cumprir o objetivo principal do nosso estudo. Além disso, relacionam-se os resultados desta investigação com estudos e teorias anteriores.

No capítulo 7, *Conclusão*, apresentam-se as conclusões gerais que se podem retirar a partir dos resultados das análises, refletindo sobre os mesmos, apresentando as limitações da investigação e sugerindo pesquisas futuras.

Capítulo 1. - A produção e a percepção da fala, e a especificidade da língua

Neste capítulo abordam-se, de forma geral, a percepção da fala e a especificidade da língua, pretendendo enquadrar a nossa investigação nas teorias da percepção da fala. Inicialmente, irá ser referido o processo da comunicação entre o falante (produção) e o ouvinte (percepção) por meio da fala, referindo-se também a categorização dos sons da fala como o processo cognitivo fundamental na sua percepção. Apresenta-se, de seguida, a especificidade da língua que altera a percepção de um indivíduo, fazendo referência à proposta de Kuhl (2004), relativa ao modo como as crianças adquirem as categorias dos sons da sua primeira língua. Por último, serão revistas as propriedades acústicas da produção e da percepção das vogais, de modo a fundamentar a escolha da teoria que guiou a nossa investigação.

1.1. Percepção da fala

1.1.1. Cadeia da fala (processo entre a produção e a percepção da fala)

Quando comunicamos com outra pessoa com a intenção de trocar ideias, utilizando a linguagem, ocorrem vários processos cognitivos (representados na Figura 1). O emprego da linguagem para trocar ideias é de tal forma comum que não notamos a complexidade dos processos cognitivos subjacentes (Fernández & Carins, 2011: 16).

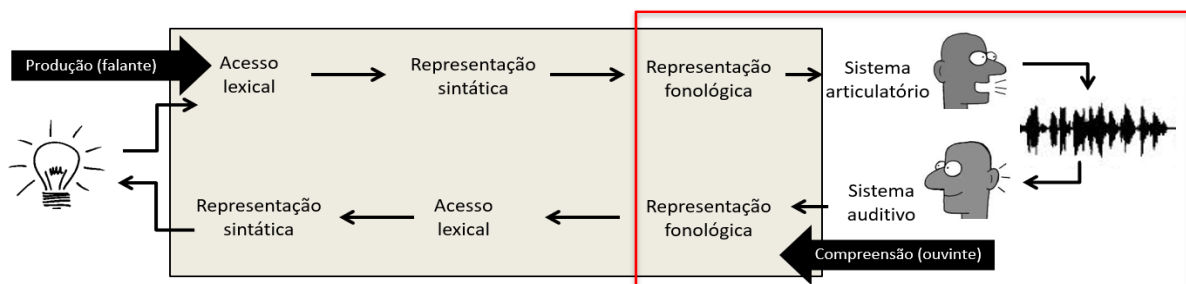


Figura 1. Processos envolvidos na produção (da esquerda para a direita) e compreensão (da direita para a esquerda) da linguagem humana. Adaptado de Fernández e Carins (2011: 16) O quadro a vermelho marca os domínios tratados nesta Dissertação.

Como se verifica na Figura 1, o falante pretende transmitir uma ideia/um pensamento ao ouvinte. De acordo com Fernández e Carins (2011), o primeiro passo do falante é traduzir o seu pensamento numa representação semântica, selecionando, no seu léxico, as palavras adequadas. A seguir, utiliza a sua gramática de modo a construir a representação sintática, conforme o significado da mensagem que quer transmitir. Estas palavras têm que ser representadas com sons, ou seja, devem ter uma representação fonológica, pois vão ser pronunciadas pelo falante. Por último, a representação fonológica ativa as áreas motoras do cérebro do falante, que vai enviar instruções aos articuladores, de maneira a produzir os sons da fala. Quando os sons da fala atingem o sistema auditivo do ouvinte, inicia-se um processo de descodificação para reconstruir a mensagem do falante. Primeiro, o ouvinte deve reconstruir a representação fonológica de modo a recuperar as palavras transmitidas pelo falante e os seus significados. O ouvinte utiliza, então, o conhecimento gramatical e lexical partilhado com o falante e organiza a estrutura das palavras. Através desses processos, o ouvinte, finalmente, tem as informações suficientes para recuperar o significado básico do enunciado produzido pelo falante (Fernández & Carins 2011: 16). Estando a comunicação através da fala assente numa cadeia de processos/eventos que liga o falante e o ouvinte, Denes e Pinson (1993) denominaram esta sequência de eventos como *Speech Chain* (cadeia da fala).

Na secção seguinte trataremos do processamento da categorização do ouvinte, descrevendo brevemente as características dos sons da fala e a representação fonológica.

1.1.2. Perceção da fala como categorização

Na perceção dos sons da fala, a qual faz parte da cadeia da fala entre o falante e o ouvinte, um ouvinte não linguista ou que não tem conhecimento fonético-acústico assumiria, facilmente, que um enunciado é uma sequência de sons discretos (Tatham & Morton, 2006: 4). Esta observação de um ouvinte leigo, todavia, não reflete fisicamente a realidade. De facto, os sinais acústicos da fala, quando representados na forma de onda, revelam um padrão não discreto, mas contínuo – cf. Figura 2.

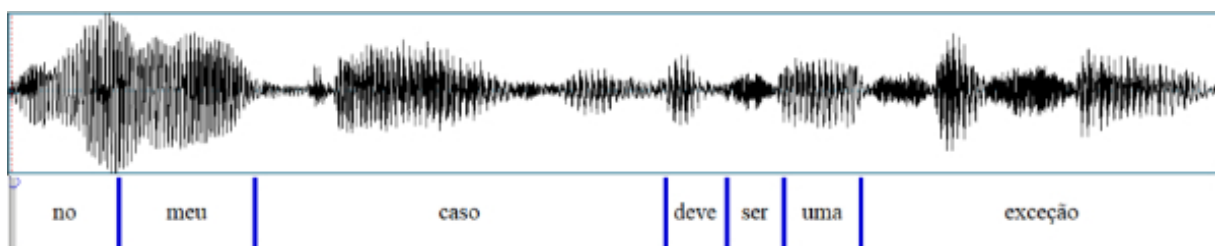


Figura 2: Forma de onda para a frase “No meu caso deve ser uma exceção”, ilustrado graficamente a natureza contínua do sinal de fala. Os limites da palavra não são particularmente salientes, nem os limites entre as consoantes e as vogais que compõem as palavras.

Além disso, estes sinais são altamente variáveis, dependentes, por exemplo, do tamanho do trato vocal, do débito da fala, da posição da sílaba, do estado emocional do falante e do ambiente em que o enunciado é produzido (Rosner & Pickering 1994, Lively & Pisoni, 1997, Holt & Lott, 2010; Fernández & Carins, 2011; Cutler, 2012). Sendo os sons da fala a única ligação física, mas não direta, entre o falante e o ouvinte, estes devem reconstruir a relação, entre o sinal físico contínuo que ele recebe e as unidades claramente estruturadas e discretas da representação fonológica, de modo a que o processo mental avance para recuperar o significado transmitido pelo falante (Fernández & Carins, 2011: 18). Portanto, apesar de não estarem conscientes disso, os ouvintes têm que determinar os limites nesses espaços acústicos para perceber quais os fonemas que o falante pretende emitir. Devido a este processo cognitivo, designado por categorização, temos a sensação de que os sons da fala são discretos (Tatham & Morton, 2006: 4). Assim sendo, muitos investigadores da percepção da fala assumem que o domínio da representação fonológica é uma etapa indispensável na compreensão da fala de uma dada língua (Holt & Lott, 2010: 1218; Eerola *et al.*, 2014: 3) e têm-se focado no mapeamento dos sinais acústicos no sistema fonológico (Holt & Lott, 2010: 1218), procurando determinar não só as pistas acústicas que contribuem fortemente para a decisão do ouvinte na categorização fonémica, como também as pistas fracas ou não pistas (*cue weighting*) (Newman, 2003: 2850). Embora este processo automático nos pareça simples, o mecanismo é na verdade complexo, não existindo ainda um modelo explicativo satisfatório que conecte o objeto físico-acústico com a representação linguística abstrata na mente humana (Iverson & Kuhl, 2000; Holt & Lott, 2010), dado que este processo não é uma deteção simples de correspondência entre os sinais acústicos e os fonemas. Holt & Lott (2010) não consideram o processo da percepção da fala nem categórico, nem mera identificação, mas sim uma categorização perceptual que se liga

diretamente a outras investigações da área da cognição humana. As características particulares deste processo que o tornam distintivo em relação a outros tipos de categorização, como a visual, são as seguintes: o espaço de *input* da fala é composto por dimensões acústicas maioritariamente contínuas, o que obriga a uma análise baseada em categorias; não há uma única pista que seja necessária ou suficiente para definir a relação entre o *input* e a categoria; sendo a fala um processo, por si só, temporal, não se pode comparar um exemplo com outro paralelamente; a informação relativa à categorização está distribuída ao longo do tempo, criando, deste modo, o problema da segmentação. Holt e Lott (2010) defendem que este processo (categorização dos sons da fala) de extração de uma regularidade através da fonte de informações acústicas ricas e variáveis merece ser colocado numa posição central nas ciências cognitivas.

Salienta-se que vários estudos da percepção de fala têm mostrado a influência da língua nativa do ouvinte na percepção, isto é, o ouvinte categoriza os sons de modo diferente, dependendo da sua língua nativa (Holt & Lott, 2010; Cutler, 2012), como abordaremos a seguir.

1.1.3. Especificidade da língua

Observando os sons das línguas do mundo, que incluem cerca de 600 consoantes e 200 vogais (Ladefoged, 2005), verificam-se algumas características comuns entre estas. Considera-se que os sons da língua são condicionados por três questões, a saber: fisiológicas (Cutler, 2012: 5), auditivas (Johnson, 2010: 119-120) e também fonológicas (Johnson, 2000: 1; Hume & Johnson, 2001: 7). Estas características universais devem-se à função central da fala, que é a comunicação: “*we speak to be heard, in order to be understood*” (Jakobson, Fant & Halle 1952:13).

Por outro lado, cada língua possui, no seu sistema, um número distinto de segmentos fonológicos que distinguem o significado das palavras, denominados (pelos estruturalistas) fonemas³. Os fonemas são as unidades abstratas cujas realizações concretas são grupos de sons não idênticos que funcionam equivalentemente numa dada língua (Kuhl, 2004; Veloso, 2015),

³ Para Troubetzkoy (1969 [1939]) e os seus seguidores estruturalistas, o fonema é a unidade linguística mínima que tem uma função distintiva numa dada língua. Por outro lado, no modelo SPE (Chomsky & Halle, 1968) a unidade mínima do sistema linguístico é definida como o traço, cuja função também é distintiva, não entre significados, mas entre segmentos. Deste modo, de acordo com este modelo, o segmento é um conjunto de traços. Tendo em conta a natureza desta investigação, estes dois termos “fonema” e “segmento” são tratados como equivalentes, neste estudo.

isto é, a troca de um som por outro não muda o sentido da palavra. De acordo com a base de dados UPSID -*The UCLA Phonological Segment Inventory Database*- (Maddieson, 1984), o tamanho típico do inventário fonológico das línguas do mundo é de 20 a 37 segmentos (Maddieson, 1984 :7), observando-se, no entanto, que existem línguas que possuem apenas 11 segmentos (Rotokas e Mura) e no outro extremo, 141 segmentos (!xũ). Esta variedade do tamanho do inventário fonológico tem implicações na percepção. A percepção da fala de um indivíduo, depende, pois, da sua língua nativa, ou seja, a sua percepção é moldada pela sua experiência linguística (Rosner & Pickering, 1994; Kuhl & Iverson, 1995; Cutler, 2012). O ouvinte categoriza os sons da fala de modo diferente, conforme o sistema da sua língua nativa (Polka, 1991; Best & Strange, 1992; Cutler, 2012). Uma vez adquirido o sistema fonológico de uma dada língua, um adulto pode ter dificuldade em distinguir o contraste dos sons não existentes na sua língua (Best & Strange, 1992). Este comportamento tem sido observado em vários estudos experimentais da aquisição da segunda língua. Um exemplo frequentemente referido na literatura será a dificuldade que os japoneses têm na distinção entre as consoantes líquidas [r] e [l], que são funcionalmente equivalentes no sistema da sua primeira língua (Miyawaki *et al.*, 1975; MacKain, Best & Strange, 1981; Mochizuki, 1981). A inexistência do contraste entre as consoantes líquidas faz com que os ouvintes japoneses não possuam sensibilidade suficiente para distinguir a diferença entre estes dois sons. Este tipo de fenómeno verifica-se também com os nativos do inglês nos contrastes de hindi entre a consoante surda aspirada vs. *breathy voiced* /t^h-/d^h/, e entre a consoante retroflexa vs. dental: /ʈ/-/ʈ/, que não são, no seu sistema, fonemas distintos (Werker & Teses, 1984b). A influência do sistema da primeira língua não ocorre apenas com os contrastes fonémicos. Por exemplo, o sistema da língua japonesa possui alguns contrastes de *mora*: o contraste entre as vogais curtas e longas e ainda entre as vogais com a mesma qualidade seguidas de consoante dupla. As diferentes *mora* (*mora* especiais) têm função distintiva. Para muitos estudantes ocidentais, cujo sistema linguístico não possui esta distinção, torna-se um desafio perceber estas diferenças (Toda, 2003).

Outra evidência da especificidade da língua resulta de uma experiência de identificação das vogais (Stevens *et al.*, 1969) de duas séries de contínuos: /i/-/ɪ/-/ɛ/ e /i/-/y/-/u/, com ouvintes americanos e suecos, sendo a primeira série comum para ambas as línguas e a segunda ocorrendo apenas em sueco. Enquanto na identificação da primeira série, os dois grupos responderam de modo semelhante, na segunda série os americanos não foram bem-sucedidos em comparação com o grupo sueco. Além disso, os limites das categorias foram marcados de

modo diferente pelos dois grupos. No estudo de Masapollo *et al.* (2017), investigou-se a assimetria da direção na percepção⁴, isto é, a discriminação é mais fácil quando os estímulos são apresentados numa determinada sequência, do que na sequência inversa, ainda que seja o mesmo par com dois estímulos. Com o objetivo de esclarecer se esta assimetria resulta dos efeitos da focalização dos formantes (um efeito universal) ou do protótipo (um efeito relacionado com uma língua específica) Masapollo *et al.* (2017) testaram, primeiro com os falantes de francês canadiano e de inglês canadiano, a localização acústica dos melhores exemplos da vogal /u/, vogal que ambas as línguas possuem, classificando com uma pontuação de *goodness* os estímulos sintetizados. O resultado mostrou que, em comparação com a classificação dos falantes nativos de inglês, os falantes nativos de francês classificaram como os melhores exemplares os estímulos mais focais, isto é, os sinais acústicos com proeminências espectrais devido à convergência espectral de alguns formantes, exibindo, assim, uma estabilidade perceptiva (Masapollo *et al.* 2017: 2859). Por outras palavras, os falantes nativos de francês escolheram, como os melhores exemplos, os estímulos de /u/ que possuíam a convergência de F₁ e F₂ com valores mais baixos do que os escolhidos pelos falantes de inglês.

Por outro lado, um curto período de experiência linguística pode também afetar a percepção de um indivíduo. Evans & Iverson (2004, 2007) investigaram o ajustamento perceptivo dos jovens universitários ao sotaque do ambiente em que estavam integrados. Verificou-se, então, uma evolução do seu sistema vocálico ao longo do tempo. No estudo da aquisição de L2, Bradlow *et al.* (1997), com o treino perceptivo da identificação das consoantes líquidas /r/ e /l/, obtiveram resultados indicadores de progresso tanto na percepção como na produção com os adultos japoneses. Consideraram-se esses resultados como uma transferência da capacidade perceptiva ao domínio da produção. Salienta-se que, no entanto, ao contrário daquilo que se pensa intuitivamente, numa ligação estreita entre a produção e a percepção da fala, nem nos estudos de L1, nem de L2 têm havido resultados consistentes, ou seja, não existem ainda evidências coerentes para provar esta relação (Evans & Iverson 2007: 3814-3815).

Na secção a seguir, veremos como as crianças aprendem a língua num determinado ambiente, ou seja, como a percepção das crianças é moldada pela língua.

⁴ A questão da universalidade entre as línguas do mundo tem suscitado grande interesse (Cutler, 2012). No estudo referido (Masapollo *et al.* 2017) observou-se, com os resultados da tarefa de discriminação obtidos, a influência da vogal focal na assimetria direcional, mas não da vogal prototípica. Os autores, contudo, salientam que estes resultados, que mostram a influência do fator universal, não excluem o efeito pela língua específica na assimetria direcional.

1.1.4. Aquisição de categorias e a alteração da capacidade perceptiva

Segundo Kuhl (2004), as primeiras tarefas das crianças na aquisição da língua serão “fazer algum progresso em descobrir as categorias fonémicas da sua língua, antes de aprender as palavras” e ao mesmo tempo “aprender a agrupar em categorias os sons que lhes são perceptivamente distintos” (Kuhl 2004: 832). Os bebés, nos primeiros 6 meses, são capazes de distinguir quase todos os contrastes fonéticos usados em línguas naturais, mesmo que estes não estejam presentes na sua língua (Werker & Tees, 1984a; Werker & Lalonde, 1988; Kuhl, 2004). Esta capacidade, todavia, vai declinando à medida que eles são expostos a uma língua específica (Werker & Tees, 1984a; Werker & Lalonde, 1988; Tsushima *et al.* 1994; Kuhl, 2004; Kuhl *et al.*, 2005), ou seja, a sua perceção vai-se alterando em função da estrutura da sua língua nativa (Werker & Tees, 1984a; Kuhl *et al.*, 2008). Tsushima *et al.* (1994) observaram que os bebés japoneses com 10-12 meses já não conseguiam discriminar o contraste /r/-/l/, enquanto que os bebés japoneses com 6-8 meses ainda o conseguiam. Concluíram que, deste modo, o processo de declínio (*developmental decline*) na discriminação entre as líquidas ocorre no final do primeiro ano⁵. Pode formular-se a hipótese que, para aprender e desenvolver melhor a sua língua nativa, os bebés “fecham” a sua capacidade perceptiva inata relativa a outros contrastes que não se ajustem à sua língua nativa (Kuhl *et al.*, 2005: 251). Esta hipótese é sustentada pelo facto de se ter verificado uma correlação negativa entre este declínio perceptivo dos bebés e o desenvolvimento posterior da linguagem. O mecanismo das duas tarefas supramencionadas é extremamente complexo, pois, por um lado, as crianças têm que adquirir as regras fonológicas que lhes indicam quais as unidades fonéticas que compõem uma categoria fonémica, e por outro lado, delimitar os sons físicos variáveis dependentes do falante, do débito de fala ou do contexto, em categorias fonémicas conforme o sistema da sua língua. Kuhl (2004) propôs um modelo para a estratégia utilizada pelas crianças, relativa à aprendizagem das categorias. De acordo com este modelo hipotético, as crianças analisam a língua utilizada pelas pessoas que a rodeiam, registando as frequências de distribuição dos segmentos fonéticos da sua língua. Por exemplo, na língua japonesa existem apenas 5 vogais fonémicas ao passo que no inglês possuem 10 e no sueco 16. Este facto, no entanto, não significa que os japoneses não produzam acusticamente uma ampla gama de vogais. O que difere entre as línguas é o padrão da distribuição dos sons,

⁵ O processo de declínio (*a developmental decline*) na discriminação de contrastes não nativos ocorre por volta dos seus 10-12 meses para algumas consoantes (oclusivas e fricativas), e pelos 6 meses para as vogais (Polka & Werker, 1994).

que fornece as pistas sobre a estrutura fonémica de uma língua (Kuhl, 2004: 833). Analisando o valor modal de cada categoria, isto é, o valor mais frequente que as crianças ouvem, agrupam todas as realizações fonéticas perto desse valor numa categoria e delimitam entre categorias onde a frequência é baixa (Kuhl, 2004: 833). A sensibilidade das crianças ao padrão distribucional da sua língua foi verificada na experiência perceptual de Kuhl *et al.* (1992). Nessa experiência, os bebês americanos e suecos agruparam apenas as variantes das vogais da sua língua, demonstrando uma distorção perceptiva influenciada pela vogal prototípica da sua língua. Esta distorção perceptiva é chamada “*perceptual magnet effect*”, isto é, a sensibilidade aos estímulos é alta quando estes são maus exemplares e, de forma oposta, quando são os melhores exemplares a sensibilidade torna-se mais baixa do que os anteriores. “It is as if the perceptual space underlying speech sounds is distorted, being stretched at category boundaries and shrunk near best exemplars” (Iverson & Kuhl, 2000: 874). A autora salienta que este efeito reflete a aprendizagem baseada nos protótipos. Voltaremos à teoria baseada nos protótipos na secção 1.3. *Teoria do protótipo: estrutura interna da categoria dos sons da fala.*

1.2. Produção e percepção das vogais

1.2.1. Produção das vogais

Do ponto de vista da fonética articulatória, os sons da fala são produzidos pelo aparelho fonador que inclui três partes: cavidades subglotais, laringe e cavidades supraglotais. Estas três partes são, por sua vez, constituídas por órgãos: as cavidades subglotais pelos pulmões, a laringe pelas pregas vocais e pela glote, e as cavidades supraglotais pela faringe, pelo trato vocal e pelo trato nasal. A fonte de energia da produção da fala é o fluxo do ar egressivo que vem dos pulmões e/ou ingressivo⁶. As vogais são produzidas quando, na laringe, as pregas vocais se aproximam, diminuindo o espaço através do qual o ar passa e, conseqüentemente, produz uma vibração em toda a extensão, sem turbulência nem nenhum tipo de obstáculo parcial no trato vocal. Para emitir uma qualidade específica da vogal, articulam-se o véu palatino, a língua, a cavidade bucal e os lábios. A posição do véu palatino define os sons como orais ou nasais, ou seja, à passagem do ar, se o véu palatino estiver levantado contra a parede posterior da faringe,

⁶ Existem línguas que possuem no seu sistema linguístico sons produzidos com o fluxo do ar ingressivo, por exemplo, os cliques (Falé, 2005: 68).

o fluxo do ar passa apenas pela via bucal, sendo emitido um som oral. Se não estiver levantado, passa tanto pela via bucal como pela via nasal simultaneamente, sendo emitido um som nasal. Para a descrição das vogais, tem-se em conta a relação⁷ entre a articulação e a qualidade da vogal. Tradicionalmente a posição da língua e dos lábios é utilizada para classificar as vogais. Em relação à posição da língua, consideram-se a posição do dorso, que se refere à altura da língua (alta, média e baixa), e a posição da raiz da língua, que se refere ao ponto de articulação (anterior/palatal, central e posterior/velar) relativamente à posição neutra, isto é, no seu estado de repouso. A respeito da posição dos lábios, descrevem-se como arredondados ou não arredondados. De acordo com os critérios articulatórios tradicionais, existem nove vogais orais fonéticas do PE padrão – cf. Tabela 1.

Tabela 1: Classificação das vogais orais do Português Europeu padrão. Adaptado de Falé (2005: 79).

	Anterior/Palatal	Central	Posterior/Velar
Alta	[i]	[ɨ]	[u]
Média	[e]	[ɐ]	[o]
Baixa	[ɛ]	[a]	[ɔ]

1.2.2. As propriedades acústicas e a percepção das vogais

Tal como referido acima, as vogais são produzidas através da vibração das pregas vocais com ausência de obstáculos no trato vocal. Estas características articulatórias das vogais têm consequências acústicas. A primeira característica faz com que as vogais sejam muito ressoantes e a segunda permite a variação temporal entre longas e curtas, ao passo que as consoantes, os sons produzidos com obstáculos, podem variar apenas dentro de um limite muito mais restrito. Considerando a ideia de que na percepção existe uma porção acústica do sinal com mais relevância face às restantes, Cutler (2012) afirma que, para as vogais, esta porção mais relevante é o estado estacionário, enquanto que para as consoantes pode ser a transição de um estado para outro⁸. Estas informações acústicas poderão ter um papel importante na percepção

⁷ Ladefoged & Johnson (2011: 87-89) refere que a descrição tradicional articulatória das vogais não é uma descrição absoluta. “But it is important for you to remember that the terms we are using are simply labels that describe how vowels sound in relation to one another. They are not absolute descriptions of the position of the body of the tongue.”

⁸ Hillenbrand (2013) defende a importância da mudança espectral inerente das vogais na percepção.

das vogais na fala. Por exemplo, por vezes as vogais são percebidas de forma mais precisa do que as consoantes num ambiente barulhento (Cutler *et al*, 2004), sendo mais provável que os erros auditivos sejam atribuídos às consoantes mal percebidas do que às vogais (Bond, 1999: 25). Cutler (2012: 6) argumenta que “the longer a sound, and the more steady-state components it comprises, the more resistant it is to noise masking”.

Por outro lado, a configuração da língua, da boca e do véu palatino determina a qualidade da vogal produzida conforme as suas posições, intensificando e/ou enfraquecendo as partes específicas do sinal acústico. Deste modo, cada vogal demonstra diferentes padrões no espectrograma (Falé, 2005: 125- cf. Figura 3). Do ponto de vista fonético-acústico, a configuração do trato vocal, mais especificamente a altura e o recuo da língua, relaciona-se com os valores físicos dos formantes: o primeiro formante (F_1) reflete a altura da língua e o segundo (F_2) o recuo da língua.

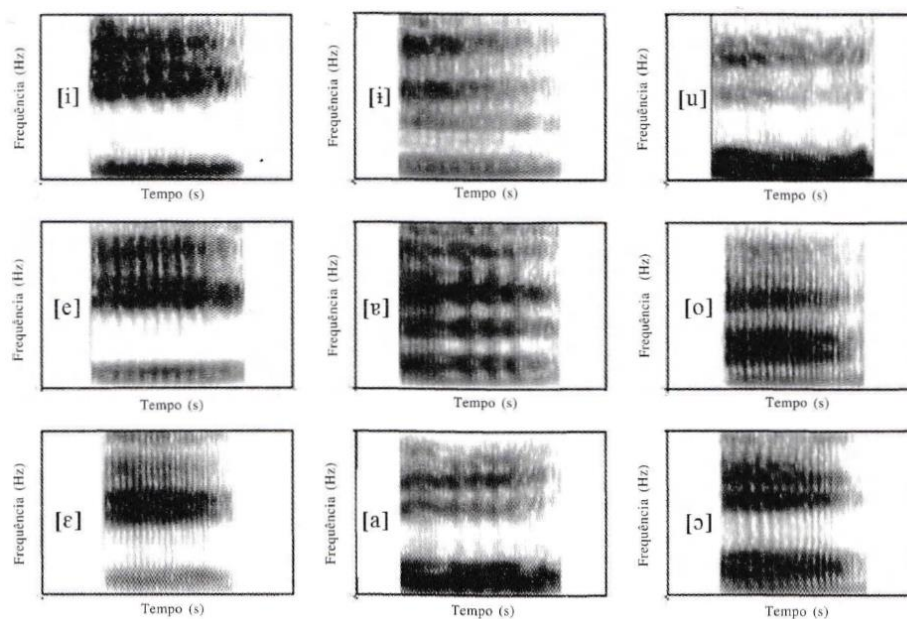


Figura 3: Espectrogramas de banda larga das nove vogais orais do português europeu padrão, produzidas por uma falante feminina. Adaptados de Falé (2005:125)

O facto de que cada vogal exhibe um padrão acústico particular ajudaria a nossa percepção na identificação da vogal que foi produzida. Embora haja um maior número de formantes, as experiências acústicas da percepção da fala têm revelado que os primeiros três formantes são mais do que suficientes na identificação das vogais. Contudo, quando um falante produz uma vogal, os valores dos formantes, de facto, apresentam dispersão e sobreposição, ou seja, estes

três por si só não permitem ao ouvinte identificar totalmente a vogal produzida (Denes & Pinson, 1993:166).

Os paradigmas distintos da percepção das vogais e das consoantes foram descobertos numa experiência da percepção da fala com a tarefa de categorização. O paradigma das consoantes foi primeiro observado por Liberman *et al.* (1957), que estudaram o efeito da transição do segundo formante na percepção das consoantes oclusivas. Confirmou-se, então, o paradigma da percepção da fala designado por percepção categorial. Isto é, numa tarefa que envolva decidir qual o som ouvido ao longo de um contínuo sonoro manipulado do valor típico de um som para o valor típico de outro som, o ouvinte é (quase) insensível à mudança dentro de uma categoria, embora essa mudança tenha sido apresentada acusticamente. Os ouvintes mostram dificuldades ao discernir a categoria apenas dentro de um espaço acústico estreito – denominada fronteira entre as categorias. Este padrão, chamado “efeito de limite de fonema” revela que, ao longo do contínuo, apesar de alguns sons serem ambíguos e nunca terem sido ouvidos na sua experiência linguística, estes tendem a ser categorizados como um membro de uma ou outra categoria de fonemas da sua língua (Cutler, 2012: 8). Partindo da hipótese de que os ouvintes conseguem discriminar os sons da fala se e só se esses pertencerem a duas categorias diferentes, Liberman *et al.* (1957) estabeleceram uma correlação entre o desempenho da identificação e o da discriminação. Isto é, o desempenho da tarefa de identificação serve para prever até que ponto os ouvintes seriam capazes de discriminar entre dois sons. Assim, o modelo da percepção categorial implica o mecanismo hipotético no qual a fala é percebida no sentido dos seus rótulos fonémicos, isto é, categorias discretas fonémicas, e não das propriedades acústicas contínuas (Iverson & Kuhl, 2000: 874). O modelo, conseqüentemente, requer que todos os membros da mesma categoria sejam percebidos como idênticos (Iverson & Kuhl, 2000: 876).

Por outro lado, em Pisoni (1973), numa experiência da mesma natureza entre as vogais /i/-/ɪ/, os resultados da identificação e da discriminação revelaram que os participantes perceberam pequenas diferenças mesmo dentro de uma só categoria, mostrando uma maior dificuldade no ato da categorização das vogais, em comparação com a de consoantes oclusivas – cf. Figura 4. Ou seja, os ouvintes revelaram a sua capacidade de captar mais informação nas vogais além dos rótulos fonémicos discretos (Iverson & Kuhl, 2000: 876). Este fenómeno, de facto, deve-se à natureza rica das vogais em termos de informações acústicas que permite ao ouvinte uma análise mais detalhada, ficando, assim, mais sensível a pequenas mudanças

acústicas (Cutler, 2012: 9). Por outras palavras, os ouvintes são mais sensíveis à estrutura interna das categorias vocálicas⁹.

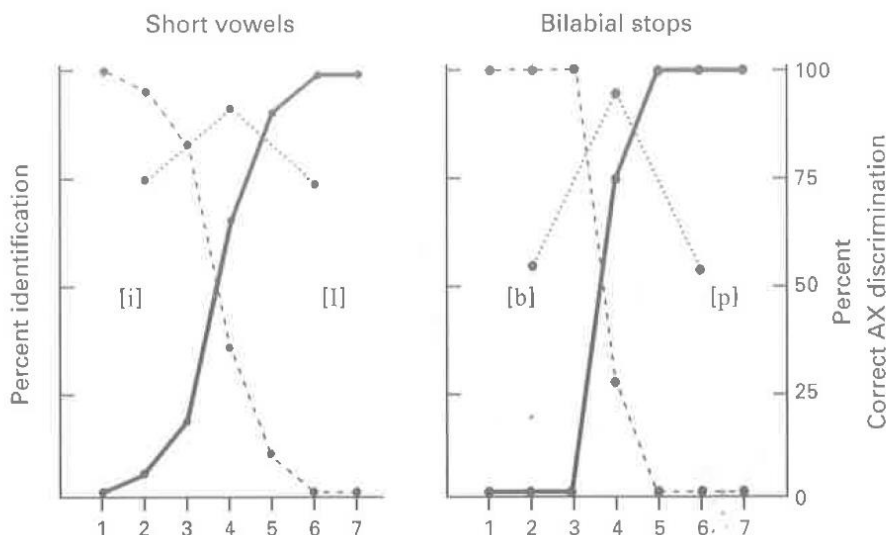


Figura 4: Representações gráficas dos dados das tarefas de identificação e de discriminação das categorias das vogais [i]-[ɪ] e das consoantes [b]-[p] (Pisoni, 1973). As linhas cruzadas representam as repostas de identificação de cada categoria e a linha pontilhada acima do cruzamento representa as repostas de discriminação. Adaptado de reprodução por Cutler (2012: 9).

1.3. Teoria do protótipo: estrutura interna da categoria dos sons da fala

A teoria do protótipo, originalmente proposta por Rosch (1975), sugere a hipótese de que todas as categorias do mundo são captadas perceptivamente de modo gradual, sendo o protótipo o seu centroide e definido como o seu melhor membro representativo. Postula, assim, a ideia de uma estrutura interna de categoria em que os membros pertencentes não possuem a mesma importância representativa. Embora a autora posteriormente mencione que os princípios que propôs devem ter em conta as questões da categorização relacionadas com uma dada cultura, codificadas pela sua linguagem num determinado tempo; e não se destinam a constituir uma teoria do desenvolvimento das categorias nas crianças nascidas numa cultura, nem a modelar o processo mental de categorização de um falante adulto de uma linguagem (Rosch, 1978: 28), a sua teoria atrai muitos investigadores na área da fonética (Lively & Pisoni, 1997). Considerando

⁹ Existe também um estudo no qual se trata da estrutura interna das categorias consonânticas. Em Iverson & Kuhl (1996) investigaram-se as categorias das consoantes líquidas /r/-/l/ do inglês americano.

as categorias fonéticas contínuas, estes investigadores aplicam esta teoria de modo a dar uma resposta ao fenómeno da categorização dos sons da fala (Oden & Massaro, 1978; Samuel, 1982; Grieser & Kuhl, 1989; Miller & Volaitis, 1989; Kuhl, 1991; Kuhl *et al.*, 1992; Lively & Pisoni, 1997). De acordo com Lively e Pisoni (1997), os fonemas prototípicos armazenados na memória a longo prazo podem servir como pontos de referência cognitivos contra os quais itens reais são julgados, isto é, os ouvintes comparam os sinais acústicos de *input* com os fonemas idealizados no seu conhecimento linguístico. Sendo um *input* semelhante ao protótipo, o melhor representante de uma dada categoria fonémica, é reconhecido pelo ouvinte como membro dessa categoria (Lively & Pisoni, 1997; Eerola, *et.al*, 2012).

Assim sendo, o protótipo tem um papel importante na estruturação das categorias da fala. Aceitando esta visão da teoria do protótipo, as fronteiras entre categorias tornam-se num epifenómeno da categorização (Samuel, 1982: 307). Este modelo, pois, implica um mecanismo hipotético no qual as categorias fonémicas são representadas em termos de protótipos como únicos representantes de todos os membros da mesma categoria e, através da distância em relação a esses protótipos, os membros são percebidos, pelo menos parcialmente, como uma categoria fonémica (Iverson & Kuhl, 2000: 874). Resultados anteriores indicam que o fenómeno da percepção categorial é observado não só nos seres humanos, mas também noutras espécies como a chinchila (Kuhl & Miller, 1975). No entanto, o *perceptual magnet effect* - a capacidade dos ouvintes discriminarem pequenas mudanças de sons não protótipos, juntamente com a sua insensibilidade à mudança dos exemplos próximos dos protótipos da mesma categoria (Kuhl, 1991; Kuhl, 1992 *et al.*; Kuhl & Iverson, 1995) - é observado apenas nos seres humanos (tanto nos bebés como nos adultos). Kuhl (1991) e Kuhl *et al.* (1992) sugeriram então que a sensibilidade desigual conforme o protótipo se deve à experiência linguística. Esta abordagem permite, assim, explicar o modo como o ouvinte categoriza os sinais da fala altamente variável intra e entre pessoas (Lively & Pisoni, 1997).

A forma como decodificamos os sinais da fala em categorias fonémicas e o seu mecanismo ainda não estão definidos. Contudo, tendo em conta os estudos referidos acima, iremos recorrer à teoria do protótipo para realizar esta investigação, pois consideramos que esta é a mais adequada à categorização das vogais na percepção.

1.4. Teoria do protótipo na base de experiências da percepção da fala com *goodness rating*

Recentemente, a teoria do protótipo tem sido empregue para investigar mais amplamente a percepção da fala com base em *goodness ratings*. Por exemplo, Masapollo *et al.* (2017) utilizaram o *goodness rating* na sua primeira experiência com 13 jovens falantes nativos monolíngues de francês canadiano, oriundos do Québec, e 13 jovens do inglês canadiano, oriundos de Ontário. Os ouvintes tinham que escolher uma palavra que continha a vogal ouvida, conforme a categoria da sua língua e, posteriormente, classificar a qualidade do som ouvido numa escala de 1 a 5 conforme a vogal selecionada, sendo 1 um mau exemplo e 5 um bom exemplo da categoria. Para os falantes de inglês apresentaram-se as palavras <boo> (para a vogal /u/), <bee> (/i/), <bowl> (/o/), <bug> (/ə/), e um <X> caso não existisse essa vogal em nenhuma das palavras. Para os falantes de francês, <boue> (/u/), <bu> (/y/), <bas> (/a/) e <beau> (/o/) e opção <X> para ausência de correspondência. Confirmou-se uma alta sensibilidade às propriedades acústico-fonéticas distintas que especificam o /u/ da sua língua nativa, a vogal-alvo desta experiência, classificando distintamente estímulos com diferentes propriedades como os melhores exemplares dessa categoria. Mais especificamente, os falantes de francês escolheram, como os melhores exemplos, os estímulos mais focais, isto é, os que se localizam mais na periferia do triângulo vocálico, tendo o valor de F_2 mais baixo no caso de /u/, em comparação com os melhores exemplares escolhidos pelos falantes de inglês. Revelou-se, assim, uma alteração perceptiva causada pela sua experiência linguística ao longo do tempo. Salienta-se ainda que houve uma correspondência estreita entre a estrutura perceptiva da categoria /u/ para cada língua e os valores dos formantes, F_1 e F_2 , da produção de cada língua, investigados nos estudos anteriores.

Em Newman (2003), recorrendo à classificação de *goodness*, verificaram-se as correlações entre a produção e a percepção da fala, comparando os valores acústicos médios da produção de um indivíduo e os do estímulo classificado com a pontuação mais alta, considerado como protótipo para tal indivíduo. Participaram 25 falantes nativos de inglês na primeira experiência, e 24 na segunda. Na primeira experiência, analisou-se estatisticamente a média do *Voice Onset Time* (VOT) da produção de /p/ e o VOT dos protótipos de cada participante. Encontrou-se uma correlação significativa entre a produção e a percepção, ou seja, quem produziu /p/ com o VOT mais longo, escolheu o estímulo com o VOT mais longo como o seu

protótipo. Além disso, a correlação entre a produção e a percepção foi verificada também na comparação de limite (fronteira) das categorias /b/ e /p/, analisando estatisticamente o VOT mínimo que os participantes aceitaram como pertencente à categoria /p/ na tarefa de percepção (média = 44.7 ms) e o VOT mínimo na tarefa de produção (média = 46.7 ms). Na segunda experiência, a comparação foi feita entre a média da produção de /ʃæ/ e o estímulo considerado como protótipo de /ʃæ/ na percepção. Postulando as propriedades espectrais como a pista mais importante na distinção das fricativas, foram medidos os seguintes valores espectrais: o centroide, a assimetria e o pico da fricção, de modo a detetar qual propriedade possui mais peso na percepção. Verificou-se uma correlação significativa entre a produção e a percepção apenas com a medida do pico espectral da fricção. Assim sendo, os picos espectrais podem ser considerados como melhores pistas na discriminação da fricativa, ao invés de outras medidas acústicas. Com estes resultados, Newman (2003) propôs que as correlações entre a produção e a percepção da fala possam servir para diferenciar o peso dos sinais acústicos, os quais funcionam como pistas acústicas na percepção da fala.

Referimos, nesta secção, os dois trabalhos recentes que adotaram a classificação de *goodness*, de modo a justificar a nossa escolha metodológica na procura de evidências científicas que suportem a teoria do protótipo. Salienta-se ainda que os seus resultados e observações corroboraram as hipóteses lançadas na introdução, uma vez que nestas investigações se confirmou uma categoria interna estruturada, e sustentaram também a ideia da correlação entre a produção e a percepção, embora sejam precisas mais evidências.

Capítulo 2. – Influência do sistema fonológico e as vogais do PE

Neste capítulo, de modo a contextualizar o presente estudo, apresentaremos conceitos básicos de fonética e fonologia, e discutiremos a influência do sistema fonológico na percepção, mais especificamente o impacto evidenciado na percepção dos sons da língua nativa pelas relações fonológicas estabelecidas no sistema. Em seguida, referimos algumas propostas relativas ao inventário vocálico fonológico do PE, tendo como objetivo problematizar a questão da nossa investigação, sustentando-a com alguns estudos do PE, nomeadamente a aquisição da vogal /i/ (Freitas, 2004), as variações fonéticas da vogal média fronteira /e/ (Barros, 1994; Rodrigues, 2003) e as características fonético-acústicas das vogais médias anteriores (Escudero *et al.*, 2009).

2.1. Distinção entre o nível fonológico e fonético

Uma clara separação entre o estudo fonológico e fonético foi proposta, por Troubetzkoy (1969 [1939]) na introdução de *Grundzuge der Phonologie*. Troubetzkoy (1969 [1939]) definiu o estudo da fonética como a disciplina que trata os sons da fala transmitidos pelo aparelho fonador, sujeitos a uma variação infinita, enquanto que a fonologia aborda os sons, em número finito, do sistema linguístico, que diferenciam os significados das palavras. O nível fonético é meramente substancial, físico, audível, podendo ser estudado com a abordagem de uma ciência natural, usando metodologias laboratoriais e experimentais (Troubetzkoy (1969 [1939]: 9). O nível fonológico, em que se encontram as unidades funcionais e os fenómenos gramaticais, é, por sua vez, teórico e abstrato, não podendo ser tratado com a mesma metodologia da fonética, mas sim com a do estudo do sistema gramatical da língua (Troubetzkoy (1969 [1939]: 11). Assim sendo, Troubetzkoy (1969 [1939]) determinou, como objetivos do estudo fonológico, constatar: 1) os sons que diferenciam o sentido das unidades linguísticas encontradas no léxico de uma dada língua, denominando estes elementos discriminativos como fonemas; 2) as regras de combinação entre estes; e ainda 3) as suas propriedades funcionais no sistema (Troubetzkoy 1969 [1939]:10-11). O autor propõe, como metodologia para detetar os fonemas existentes no sistema linguístico de uma dada língua, encontrar pares mínimos. Isto é, um par de palavras existentes no léxico que possuam o mesmo número e a mesma ordem de segmentos sonoros, salvo um que diferencia o significado entre elas (1969 [1939]: 48-49).

Veloso (2015) resumiu a tarefa do fonólogo como uma observação “para além da realidade puramente fonética, em busca de regularidades de natureza mais abstrata, inacessíveis à pura observação empírica dos factos acústicos” (Veloso, 2015: 6), postulando a existência de um nível abstrato “que só é alcançável a partir da análise racional e do encadeamento lógico-argumentativo” (Veloso, 2015: 11), na medida em que os objetos fonológicos “têm um carácter eminentemente (ou mesmo obrigatoriamente) conjectural” (Veloso, 2015: 11).

2.2. Influência do sistema fonológico na percepção da fala

Embora haja distinção entre o nível fonético e o nível fonológico, existem estudos que visam relacionar os dois níveis. Estes averiguam a interação entre a percepção da fala e o sistema fonológico de modo a compreender, de forma mais global, o sistema dos sons da língua. Para tal compreensão, é, portanto, necessário procurar uma solução para o problema clássico da interface entre fonética e fonologia, dado que o domínio teórico abstrato não se liga diretamente ao mundo físico. Por exemplo, Hume & Johnson (2001) propuseram um modelo cujo domínio da representação cognitiva atua como interface entre o domínio abstrato teórico e os fatores exteriores. De acordo com este modelo, os fatores exteriores como a percepção, a produção, a generalização e a conformidade, estabelecem uma relação bilateral com o domínio da representação cognitiva, podendo, assim, influenciar indiretamente a forma teórica¹⁰. Destacando, sobretudo, o papel da percepção da fala na fonologia, os autores argumentam, com uma breve contextualização dos estudos do sistema dos sons da fala, a existência de uma correlação entre a percepção da fala e a fonologia. Segundo os mesmos, o interesse pelo papel da percepção dos sons da fala no sistema fonológico levou Jakobson, Fant & Halle (1952) a descrever as unidades fonológicas e os processos fonológicos através de características acústicas/auditivas. Esta abordagem, do ponto de vista do ouvinte, no entanto, não obteve grande sucesso, uma vez que faltavam, nessa altura, tecnologias que possibilitassem o fornecimento dos dados necessários para analisar a percepção da fala. Por outro lado, a

¹⁰ No modelo referido, as capacidades perceptivas e produtivas podem ser influenciadas pelo sistema da língua através da representação cognitiva, e podem também influenciá-lo de modo indireto. A generalização refere a tendência da representação cognitiva de simplificar a realidade sensorial, ou seja, a tendência de categorização. Por último, a conformidade trata dos fatores sociais e comunicativos, como por exemplo, a norma e a linguagem sociocultural. Estes dois últimos fatores estabelecem também uma ligação indireta com o sistema, podendo, assim, influenciá-lo. (Hume & Johnson, 2001:13-14)

abordagem do ponto de vista do falante, isto é, a descrição articulatória dos sons da língua, tem sido amplamente adotada pelas teorias fonológicas *standard*, devido ao facto de que os movimentos dos articuladores (os lábios, a mandíbula e a língua¹¹) são observáveis e funcionam, conseqüentemente, como os pontos de referência, que permitem às teorias da fonologia explicar um conjunto de padrões fonológicos. Contudo, o progresso da tecnologia e da ciência permitiu grandes avanços no estudo da relação entre a percepção da fala e o padrão do sistema fonológico. Foi desenvolvida uma teoria - *Optimality Theory* (Prince & Smolensky 1993) que explica a interação dinâmica entre as restrições baseadas perceptivamente e as restrições motivadas por princípios gerais. Este progresso tecnológico e científico, e o desenvolvimento desta teoria têm aumentado o interesse dos fonólogos (Hume & Johnson, 2001: 5).

Uma das evidências desta correlação é o impacto que o sistema fonológico tem na percepção da fala. A interferência do sistema fonológico da língua nativa começou a ser discutida nos anos 30 pelos teóricos, por exemplo, Troubetzkoy (1969 [1939]), propondo a hipótese de que o indivíduo percebe os sons da fala de L2 através do filtro de L1 (Iverson, Wagner & Rosen, 2016: 1799). Mais recentemente, os estudos da aquisição/ aprendizagem da segunda língua têm verificado que não é apenas a presença e ausência de contraste de L2 no sistema de L1 que influencia a percepção. Best e Tyler (2007: 17) defendem que “nonnative speech perception is systematically affected by fine-grained phonetic similarities and dissimilarities between native and nonnative phones¹²”. De acordo com estes autores os fatores que afetam a percepção dos sons não nativos são: “native phonotactic biases”, “co articulatory patterns” e “allophonic or other phonetic variations” de L1 (2007: 17).

Por outro lado, mesmo na percepção da língua nativa refletem-se relações fonológicas mais subtis, isto é, o modo como os membros do mesmo sistema se relacionam tem conseqüências perceptivas (Hume & Johnson, 2003, Boomershine *et al.* 2008). É este o ponto que focamos no presente estudo. Já Troubetzkoy (1969 [1939]) mencionava esta influência das relações do sistema fonológico, na percepção. Um exemplo dado por Troubetzkoy (1969 [1939]) trata a oposição neutralizável, isto é, a oposição que se neutraliza em certos contextos. A oposição constante é, pelo contrário, uma relação de fonemas que nunca se neutraliza em

¹¹ Cf. nota de rodapé 7.

¹² Tendo em conta os resultados dos estudos experimentais da percepção da fala, a sensibilidade relativa à categoria interna fonética (Miller, 1994; Miller & Volaitis, 1989), e a assimetria dessa sensibilidade conforme a distância dos bons exemplares da categoria (Grieser & Kuhl, 1989; Iverson & Kuhl, 1995; Kuhl, 1993), Best & Tyler (2007) mencionaram que não são surpreendentes as influências perceptivas pelas (dis)semelhanças entre L1 e L2 dos aprendentes serem subtis (2007:17).

nenhum contexto. Segundo o autor, a diferença perceptiva entre a oposição constante e a oposição neutralizável é significativa, sendo a primeira percebida mais claramente do que a segunda (Troubetzkoy 1969 [1939]: 78). A suposição de Troubetzkoy (1969 [1939]) é atualmente suportada pelos resultados obtidos nos seguintes estudos experimentais (entre outros).

Hume & Johnson (2003) testaram a existência de diferença, na percepção, entre os contrastes parciais (que se neutralizam em certos contextos), e os contrastes não parciais (que não se neutralizam em nenhum contexto) existentes nos tons do mandarim, comparando falantes nativos do mandarim com os de inglês americano (língua que não possui esses contrastes). Através dos resultados obtidos, confirmou-se uma redução da sensibilidade na percepção dos contrastes parciais. Este resultado levou os autores a sugerirem a necessidade de serem consideradas relações fonológicas na criação de um modelo preditivo da influência da fonologia na percepção da fala. Em Boomershine *et al.* (2008) foram examinadas as diferenças perceptivas entre os sons alofônicos e contrastivos no inglês e no espanhol. Utilizaram-se como estímulos as consoantes [ð], [d], [r], dado que em inglês [d], [r] são os sons alofônicos do mesmo fonema, mas [ð] é sempre contrastiva; enquanto que em espanhol [ð], [d] são alofones e [r] é contrastiva. Os resultados mostraram que os nativos de cada língua têm menor sensibilidade à diferença entre os sons alofônicos do que aos sons contrastivos da sua respetiva língua.

Na secção seguinte, referimos algumas propostas relativas ao inventário vocálico fonológico do PE, de modo a argumentar a escolha do tema deste estudo: as vogais médias anteriores que estabelecem relações complexas entre a forma subjacente e a de superfície.

2.3. Propostas teóricas do inventário fonológico das vogais do PE

De acordo com Mateus & Andrade (2000: 17), para descrever o sistema vocálico do PE é necessária uma divisão em dois subsistemas: um de vogais tónicas e um de vogais átonas, devido à relação estreita entre a qualidade da vogal e o acento da palavra desta variante¹³. O segundo subsistema divide-se ainda em três subconjuntos conforme a sua posição na palavra e a posição do acento – cf. (1).

¹³ “Vowel quality is significantly more independent from word-stress in BP than in EP. This fact entails considerable difference between the two varieties at the phonetic level when it comes to the atonic sub-system.” (Mateus & Andrade, 2000: 17)

Vogais tónicas

['i] silo ['silu]

['e] selo ['selu]

['ɛ] selo ['sɛlu]

['a] bala ['balɐ]

['ɔ] bola ['bɔlɐ]

['o] bola ['bolɐ]

['u] bula ['bulɐ]

Vogais átonas não finais

Átonas pretónicas

[i] mirar [mi 'rar]

[u] morar [mu 'rar]

[u] murar [mu 'rar]

[ɐ] pagar [pɐ 'gar]

[i] pegar [pi 'gar]

Átonas pós-tónicas:

[i] dúvida ['duvidɐ]

[u] pérola ['perulɐ]

[u] báculo ['bakulu]

[ɐ] ágape ['agɐpi]

[i] cérebro ['sɛribru]

Vogais átonas finais

Átonas finais

[i] júri ['zuri]

[i] júre ['zurɨ]

[u] juro ['zuru]

[ɐ] jura ['zurɐ]

(1) *Inventário fonético das vogais orais e os exemplos do PE. Adaptados de Mateus & Andrade (2000: 17-18).*

Como verificado acima no inventário fonético, existem sete vogais orais na posição tónica e quatro na átona, em PE. Confirma-se, ainda, [ɐ]¹⁴ na posição tónica em palavras como “c[ɐ]ma”, “t[ɐ]lha”. Os autores referidos defendem que esta não é a vogal subjacente, mas a vogal alternada de [a] e [e] pelo contexto onde está inserida - seguida por uma consoante nasal ou por um segmento palatal, sendo o primeiro caso chamado de processo de assimilação e o segundo de centralização (Mateus & Andrade: 2000:19). Dado que este último é um fenómeno

¹⁴ [ɐ] na posição acentuada ocorre também nas palavras funcionais como “c[ɐ]da” e “p[ɐ]ra” (Veloso, 2016).

que envolve uma das variações da realização da vogal em questão no nosso estudo, voltaremos, mais à frente, a este tópico na secção 2.5. *Variações/ alternâncias ao nível fonético*.

Relativamente ao inventário fonológico do PE, na proposta de Mateus (2003), e outros autores, as realizações fonéticas na posição tónica correspondem às sete vogais segmentais do sistema fonológico do PE /i, e, ε, a, ə, o, u/, uma vez que se constata pares lexicais cujos significados não se relacionam. Portanto, confirmam-se as oposições entre as vogais semifechadas e semiabertas no sistema do português- cf. (2).

/E/	/O/
<i>sede</i> ['sedɨ] ~ <i>sede</i> ['sedɨ]	<i>molho</i> ['moʎu] ~ <i>molho</i> ['moʎu]
<i>pega</i> ['peɣe] ~ <i>pega</i> ['peɣe]	<i>forma</i> ['fɔɾme] ~ <i>forma</i> ['fɔɾme]
<i>besta</i> ['bestɐ] ~ <i>besta</i> ['bestɐ]	<i>cor</i> [kɔɾ] ~ <i>cor</i> [kɔɾ]

(2) *Os exemplos dos pares lexicais do português baseados na oposição {semiaberto *semifechado}. Adaptado de Veloso (2016: 644).*

Enquanto que muitas línguas românicas perderam esta oposição (vinda do contraste da quantidade do Latim entre breve e longa) o português ainda a conserva (Veloso, 2016: 644).

Na linha da teoria autosegmental, Wetzels (1992) apresentou uma proposta de descrição do vocalismo fonológico, com quatro graus de abertura. Este autor aplicou a teoria geométrica de Clements (1991) - a conceção da existência do nó independente da abertura vocálica que liga diretamente o *tier* vocálico - à análise das vogais do português do Brasil (PB). Esta descrição adapta-se também ao sistema vocálico do PE (Veloso, 2016: 645), descrevendo as oposições entre semifechadas e semiabertas como a oposição entre [aberto 2] e [aberto 3] - cf. (3).

[aberto 1]	/i, u/
[aberto 2]	/e, o/
[aberto 3]	/ε, ə/
[aberto 4]	/a/

(3) *A descrição dos quatro graus de abertura vocálica do português de acordo com um modelo escalar de traços unários. Adaptado de Veloso (2016: 645).*

Em PE, na posição átona, podem encontrar-se quatro vogais [i, i, e, u] numa posição não final¹⁵ e três [i, e, u] na final¹⁶. Em Mateus e Andrade (2000) e Mateus (2003), consideram-se estas vogais presentes nas sílabas átonas, como derivadas de um dos sete segmentos fonológicos encontrados na posição tónica. A sua qualidade vocálica é alterada através de um processo fonológico. Fazendo uma comparação morfofonológica, verifica-se a relação correspondente entre a forma subjacente e a forma de superfície (Mateus, 2003: 1010) - cf. (4).

Posição tónica → posição átona

f[i]ta → f[i]tinha

g[a]to → g[ɐ]tinho

d[e]do → d[i]dada

vir[a]r → vir[ɐ]

bat[e]r → bat[i]

p[ɔ]rta → p[u]rteira

f[ɛ]sta → f[i]stejo

f[o]go → f[u]gueira

m[ɛ]l → m[i]lado

f[u]ro → f[u]rado

(4) *Os exemplos da alternância vocálica da posição tónica para átona. Adaptado de Mateus (2003: 1011).*

/a/ fonológico realiza-se como [ɐ] fonético;

/e/ e /ɛ/ fonológicos realizam-se como [i] fonético;

/o/ e /ɔ/ fonológicos realizam-se como [u] fonético;

/i/ e /u/ não se alteram.

(Mateus, 2003: 1011)

¹⁵ Veloso (2016) chamou a atenção para as ocorrências excepcionais na posição átona pré-tónica como 1) [e]uro 'peu, aç[o] 'itar; 2) infl[a] 'ção, ec[ɔ] 'nomia; 3) [ɛ]colo 'gista, [o]bri 'gado e na pós-tónica como 4) re 'pórt[ɛ]r, `plânct[ɔ]n, sendo 1) resultado do processo pós-lexical de formação do ditongo do PE, e não de formas lexicais; 2) exceções lexicalmente marcadas; 3) e 4) resultados de regras excepcionais que impedem a redução do vocalismo átono. (2016:639)

¹⁶ Sobre as realizações das vogais átonas na posição final, Veloso (2016) menciona terminações excepcionais [i] [ɛ] nos casos como táx[i], biquín[i] e inclusiv[ɛ], exclusiv[ɛ]. Embora estas possam sofrer alteração pela regularização fonológica, a força do neologismo faz com que, atualmente, diminua a produtividade da redução vocálica do PE (2016:639).

Deste modo, as oposições entre [aberto 2] e [aberto 3] neutralizam-se no contexto átono, realizando-se como [i] para anteriores e [u] para recuadas¹⁷.

Embora esta proposta com sete vogais fonológicas seja habitualmente, e de forma mais ampla, aceite nesta área, a inventariação fonológica das vogais do português tem sido, atualmente, discutida no estudo da descrição fonológica desta língua (Veloso, 2016: 641). Por exemplo, considerando o processo da redução vocálica, que resulta na mesma vogal fonética para cada par de vogais médias, Brandão de Carvalho (2011) propôs uma única forma teórica para estes pares, /E/ e /O/. Ou seja, de acordo com o autor, a redução vocálica do PE aplica-se ao sistema de cinco vogais, onde a oposição entre vogais médias se neutraliza previamente (Carvalho, 2011:59). A inexistência das oposições entre [aberto 2] e [aberto 3] ao nível subjacente é defendida também por Wetzels (1992, 2011) no seu estudo fonémico das vogais do PB. Os seus principais argumentos são: 1) uma grande parte da distribuição das vogais médias é condicionada por contextos métricos, morfológicos ou segmentais, portanto é previsível (Wetzels 2011: 332); 2) há poucas exceções a esta distribuição; e 3) pares lexicalmente opostos são raros¹⁸. Deste modo, o sistema fonológico das vogais do português reduzir-se-ia a cinco vogais, /i, E, a, O, u/, tal como confirmado noutras línguas românicas.

Por outro lado, Veloso (2016) defende a existência simultânea de oposição lexical entre [aberto 2] e [aberto 3] e a existência de uma vogal subjacente como /E/ e /O/ no sistema fonológico do PE, explicando dois tipos de oposição desta língua: oposição lexical e oposição gramatical. As oposições lexicais referem-se a pares lexicais, que contrastam apenas pelo seu grau de abertura no nível subjacente - cf. (2), isto é, o seu grau de abertura é especificado no léxico e não por nenhuma derivação morfofonológica. Embora os exemplos não sejam abundantes, são atestados na língua, logo, /e/ e /ɛ/, /o/ e /ɔ/ opõem-se no nível lexical. Quanto às oposições gramaticais, referem-se a pares de palavras com radicais que possuem o mesmo

¹⁷ De acordo com a tipologia do processo de neutralização, proposta por Troubetzkoy (1939), Bisol & Veloso (2016) categorizaram as neutralizações das vogais médias do PE. Dos quatro tipos de neutralização de Troubetzkoy (1939), a neutralização das vogais médias anteriores do PE, segundo os autores, pertence ao tipo (1b), isto é, o som resultante do processo não é idêntico a nenhum dos elementos de neutralização, e tem a sua própria especificação; e na das recuadas (3b), o som resultante é um dos elementos mais extremos da oposição gradual (Bisol & Veloso, 2016: 70).

¹⁸ Wetzels considera os casos 2) e 3) lexicalmente marcados.

significado, ou seja, nestes pares não há contraste de significado, mas de classe gramatical - cf. (5)¹⁹.

(5) eu g[ɔ]verno vs. o g[o]verno; eu [ɛ]rro vs. o [e]rro; eu ch[ɔ]ro vs. o ch[o]ro; eu alm[ɔ]ço vs. o alm[o]ço

Estas oposições gramaticais resultam do processo da alternância morfofonológica, no qual é atribuído o grau de abertura. Para estes casos, podemos postular uma vogal menos especificada (com o grau de abertura não marcado) no nível subjacente: /E/ e /O/.

Outra proposta de Veloso (2012, 2016) é a inclusão das vogais centrais não baixas [i] e [ɐ] no sistema fonológico do PE. O autor admite que [i, ɐ] na posição átona não final são formas derivadas de /e, ɛ/ (ou /E/) e /a/, ou seja, as suas realizações são puramente fonéticas, tal como as teorias anteriores defendem. Contudo, verificam-se, no nível fonético, [i, ɐ] que não são derivados de formas subjacentes de /e, ɛ/ ou /a/, por exemplo, [i, ɐ] nas terminações nominais, cas[ɐ] e bas[i], e em clíticos, qu[i], d[i], m[i], lh[i], s[i], etc. Segundo o autor, para estes casos é legítimo postular a existência das formas teóricas /i/ e /ɐ/.

Tendo em conta: 1) a existência da oposição entre [aberto 2] e [aberto 3], postulando simultaneamente uma vogal teórica menos especificada comum para estas mesmas vogais e 2) a existência das formas teóricas das vogais centrais não baixas /i, ɐ/, Veloso (2016: 658) propõe uma descrição fonológica vocálica mais extensa e complexa (do que as outras duas teorias referidas anteriormente) em que se encontram nove vogais teóricas, /i, e, ɛ, a, o, ɔ, u, ɐ, i/, especificadas lexicalmente e relacionadas de modo multívoco com as formas de superfície. Deste modo: 1) um mesmo segmento pode realizar-se, no nível fonético, como sons distintos (6); 2) o mesmo som fonético pode ter dois segmentos diferentes no nível subjacente (7); 3) a inclusão das vogais centrais não baixas no nível subjacente, /i, ɐ/, faz com que as relações das

¹⁹ Trata-se da harmonização vocálica entre a última vogal do radical derivacional e o marcador temático verbal (1) / nominal (2):

(1) /E V O/_{UVRD}&MT_[aberto 4] (Verbo_1^aconjug: MT=/a/)→[ɛ V ɔ]_{UVRD}[aberto 3]

(2) /E V O/_{UVRD}&MT__[aberto 1](Nome: MT=/u/)→[e V o]_{UVRD}[aberto 2]

Anotação: &=combinação de raiz derivacional com marcador temática teórica foneticamente expresso ou omitido. (Veloso, 2016: 647)

realizações fonéticas das vogais átonas com as formas teóricas sejam mais complexas do que nas propostas anteriores – cf.- (8).

(6) dedo - dedada {/e/[e] V[i]}, seda- sedaço {/ɛ/[ɛ] V[i]}

(7) fogueira (fogo) – furador (furo) {[u]=o/V/u/}

(8) [ɐ] =/ɐ/; [i] =/i/; [u] =/u/ ou [ɐ] =/a/; [i] =/e, ε /; [u] =/o, ɔ/

Para desenvolver o nosso estudo, assumiremos a proposta de Veloso (2016), e concentrar-nos-emos, em diante, na relação entre a forma subjacente e forma de superfície das vogais médias anteriores.

Vimos, até aqui, as propostas relativas ao sistema vocálico do PE. A proposta de Veloso (2016) leva-nos a considerar o reflexo destas relações complexas das vogais médias anteriores na percepção da fala. De modo a desenvolver uma hipótese, abordaremos nas secções que se seguem estudos relevantes sobre a aquisição da vogal [i] do PE pelas crianças (Freitas, 2004) e as variações/alternâncias da vogal média anterior /e/ no nível fonético (Barroso, 1994; Rodrigues, 2003) e, por último, a descrição acústica das vogais do PE (Escudero *et al.*, 2009).

2.4. Aquisição do processo fonológico (neutralização) do PE pelas crianças

Um dos contributos do estudo da aquisição fonológica da primeira língua para a linguística teórica é a de testar hipóteses teóricas do sistema dos adultos alusivas ao modo “como se estruturam e funcionam as gramáticas das línguas” (Amorim, 2014: 2). Tal como referido na secção anterior, o sistema vocálico do PE é caracterizado pela sua rica variação morfofonológica. A aquisição das regras fonológicas é fundamental, dado que, por meio destas regras, as crianças aprendem quais os contrastes que são importantes no sistema da sua língua e desenvolvem a representação fonológica subjacente (Fikkert & Freitas, 2006: 84). Embora a aquisição das alternâncias morfofonológicas possa ter um papel significativo nas crianças para construir o sistema fonológico do adulto, na literatura existem poucos estudos relativos a esta questão (Fikkert & Freitas, 2006: 85). O objeto de estudo da nossa investigação são as vogais médias anteriores, e por isso destacamos a investigação da aquisição do [i] pelas crianças falantes de PE (Freitas, 2004). Uma vez que o modo como as crianças portuguesas adquirem o

som [i̯], que possui diferentes estatutos a nível teórico e a nível fonético, poder-nos-á revelar a complexidade do sistema desta língua.

De acordo com a proposta de Veloso (2016), o [i̯] fonético pode ser a realização fonética de ambos os segmentos /e, ε/ no contexto átono e, ao mesmo tempo, pode também ser a do segmento /i/ para os nomes de tema em *-e* e os clíticos. Além disso, esta vogal é conhecida como vogal epentética na posição de núcleo vazio para reconstrução de sílaba, resolvendo, deste modo, uma sequência consonântica problemática que viola os princípios da boa formação silábica²⁰. Existem dois tipos de [i̯]: [i̯] epentético e, [i̯] neutralizado. Freitas (2004), com o objetivo de descrever o comportamento desses dois tipos de [i̯] no sistema fonológico das crianças na fase do desenvolvimento do sistema gramatical, observou a produção dessa vogal em 8 crianças, monolíngues do PE com idade entre 0:10 a 3:7 anos, através dos dados adquiridos num estudo de caráter transversal e longitudinal.

Nas produções iniciais, foi confirmado o uso de [i̯] prosódico no contexto de: (1) *left-edge* das palavras; (2) *right-edge* se a última consoante for a líquida [r] ou [ʁ]; (3) estruturas silábicas problemáticas, tal como *sC-cluster* no início da palavra ou ataque ramificado, preenchendo o núcleo vazio. A autora referiu que as crianças começam a usar [i̯] prosódico bastante cedo e mantêm-no para desenvolver o seu sistema prosódico. O [i̯] neutralizado, por outro lado, não foi observado nas primeiras palavras das crianças, aparecendo, tardiamente, com as seguintes estratégias: (1) [i̯] produzido; (2) [i̯] eliminado; (3) produzido como [i]; (4) eliminação da sílaba em que [i̯] ocupa a posição do núcleo; (5) produzido como outras vogais ([e, u, e, o, ε]).

Embora neste estudo não tenha sido abordado [i̯] como a realização sem alternância da forma teórica /i/, os resultados mostram a probabilidade de as crianças distinguirem os estatutos diferentes de [i̯], e ainda, levam-nos a pensar na dificuldade de adquirir a relação complexa que esta vogal estabelece com outras vogais no mesmo sistema²¹.

²⁰ Este comportamento especial de [i̯] leva Veloso (2010, 2012, 2016) a defender que /i/ é uma vogal não marcada no sistema do PE. E ainda, sendo esta não marcada, é a vogal mais vulnerável à omissão na fala espontânea (Fikkert, 2005: 12).

²¹ Peperkamp & Dupoux (2000) argumentaram que as crianças podem aprender as regras alofónicas sem conhecimento semântico, deduzindo-as através da distribuição dos sons alofónicos. Sugeriram ainda que as crianças podem inferir as formas subjacentes dos sons alofónicos. Existem, no entanto, casos onde é impossível inferir as formas subjacentes sem este conhecimento, por exemplo, a redução vocálica da língua de Chamorro na posição átona (Peperkamp & Dupoux, 2000: 17). Por conseguinte, segundo as autoras, a aquisição destes poderá ser mais tardia. A neutralização das vogais médias pela redução vocálica do PE também é um destes casos (Fikkert & Freitas, 2006: 87).

2.5. Variações/ alternâncias no nível fonético

Como referimos na secção 2.3. *Propostas teóricas do inventário fonológico do PE*, Veloso (2016) dividiu os diferentes tipos de oposições entre nível lexical e morfofonológico. A centralização da vogal média alta (semifechada) acentuada no contexto pré-palatal é uma alternância morfofonológica frequente. Estas variações foram investigadas por Barros (1994) e, posteriormente, por Rodrigues (2003). Tendo como objetivo verificar a influência dos fatores sociolinguísticos, tal como a idade e o grau de instrução de falantes de Lisboa, Barros (1994) investigou as diferentes realizações possíveis das vogais médias anteriores /e/ e /ɛ/ em posição pré-palatal e também as mesmas vogais ditongadas /ej/ e /ɛj/. Para a vogal /e/ em posição pré-palatal, registaram-se as seguintes realizações: [e, ɐ, aj, a, aj, ɛ] e para o ditongo /ej/: [e, ej, ɐ, ɛj, a, aj, ɛ, ɛj]. Em ambas as formas, o fenómeno da centralização é o mais significativo (mais de 95 % em /e/ e 99% em /ej/), no entanto, a centralização não é um processo categórico, visto que existem outras variações, embora sejam pequenas percentagens do total das ocorrências. Relativamente às variações de /ɛ/ e /ɛj/, não se confirma a aplicação do mesmo fenómeno a estas formas, a não ser em casos esporádicos. Em Rodrigues (2003), um estudo comparativo de dois dialetos do PE (Lisboa e Braga)²², também foi verificada a relevância do fenómeno da centralização, sobretudo no dialeto de Lisboa- cf. Tabelas 2 e 3. Os dados obtidos por Rodrigues (2003) corroboram, de certo modo, os do estudo de Barros (1994).

²² As faixas etárias e o grau de habilitação académica dos informantes, em ambas as investigações referidas, são mais vastos do que a nossa investigação. No entanto, dado que o fenómeno de centralização de /e/ e as suas variações são significativos entre os falantes de Lisboa, consideramos legítima a referência a estes estudos.

Tabela 2: As realizações fonéticas de /ɛ/ no contexto antes de consoante palatal, pelos falantes do dialeto de Lisboa. Adaptado de Rodrigues (2003: 129).

Segmento à direita	[ɛ]		[e]		[ɛj]		[ɐ]		Total
	Nº	%	Nº	%	Nº	%	Nº	%	
[λ] ²³	31	100	0	0	0	0	0	0	31
[ʒ] em ataque silábico	7	77.7	0	0	1	14.5	1	14.5	9
[ʃ] em coda	29	93.5	2	6.4	0	0	0	0	31
Total	67	94.3	2	2.8	1	1.4	1	1.4	71

Tabela 3: As realizações fonéticas de /e/ no contexto antes de consoante palatal, pelos falantes do dialeto de Lisboa. Adaptado de Rodrigues (2003: 130).

Segmento à direita	[ɐ]		[ɛj]		[e]		[ej]		[ɛ]		Total
	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	
/ɲ/ ou /ʎ/	243	93.8	0	0	16	6.1	0	0	0	0	259
/ʃ/ ou /z/	44	64.7	22	32.3	0	0	1	1.4	1	1.4	68
/S/	10	6.7	33	22.1	101	67.7	0	0	5	3.3	149
Total	297	62.3	55	11.5	117	24.5	1	0.2	6	1.2	476

Tendo em consideração o processo de centralização, em pares como “esp[ɐ]lho” (nome) e “esp[a]lho” (verbo, 1ª pessoa do singular do presente do indicativo <espalhar>) não há contraste fonémico entre /ɐ/ e /a/, mas sim entre /e/ e /a/ ao nível morfofonológico, sendo a

²³ Não existem palavras com /ɛɲ/ (Rodrigues, 2003: 129)

forma subjacente /e/ especificada de forma diferente na realização fonética²⁴. Assim sendo, a equação (6) tornar-se-á em:

(9) /e/=[e]V[i]V[ɐ]

Esta alternância possuiu o potencial de causar mais um “ruído perceptivo” na decisão da categorização das vogais médias, isto é, uma ambiguidade perceptiva proposta por Troubetzkoy (1969 [1939]) no processamento da percepção da fala. Além disso, as frequentes variações de /e/ levam-nos a pensar numa hipótese da entidade abstrata /e/ ser o membro menos “consistente” do sistema, dado que este é construído pelas relações de oposição entre os membros.

2.6. Características fonético-acústicas das vogais médias anteriores de Lisboa

Tal como referido no ponto 1.1.3. *Especificidade da língua*, a nossa percepção é alterada pela exposição à língua do ambiente onde o indivíduo se insere. Pode, assim, supor-se que a produção de uma dada língua se reflete na percepção e vice-versa, embora ainda não existam evidências coerentes que assegurem esta suposição.

O estudo acústico mais recente e abrangente, relativo à descrição da produção das vogais do PE, foi o de Escudero *et al.* (2009) em que extraíram cinco parâmetros (f_0 , F_1 , F_2 , F_3 e duração) referentes a sete vogais tónicas [i, e, ε, a, ɔ, o, u] do PE e do PB, produzidas no contexto CVCV por 20 jovens adultos de Lisboa (sexo feminino =10) e 20 de São Paulo (sexo feminino =10). Embora se verifique dispersão da produção intra-falante e inter-falante relativamente aos valores de F_1 e F_2 , os resultados mostram o triângulo vocálico desta língua e revelam quatro zonas de F_1 (que reflete a altura da língua), onde se agrupam as sete vogais (Escudero *et al.*, 2009: 1389). Esta última constatação coincide com a descrição fonológica de Wetzels (1992) e Veloso (2016). Salienta-se ainda, no trabalho de Escudero *et al.* (2009), uma

²⁴ Fikkert e Freitas (2006) investigaram a aquisição da alternância morfofonológica das vogais subjacentes /a/ e /e/ no contexto pré-palatal cuja realização fonética é a mesma: [ɐ]. As variações e os erros produzidos pelas crianças do PE, sobretudo com a forma subjacente /e/ no contexto referido, revelam uma aquisição tardia das regras das alternâncias morfofonológicas. As autoras, no entanto, acrescentaram que este processo começa mais cedo para as crianças falantes do PE do que das línguas germânicas, presumindo que este facto pode ser devido à complexidade do sistema vocálico do PE que envolve muitos processos, tal como a redução vocálica, a harmonização vocálica, supressão da vogal e variações alofónicas.

análise de dois dialetos do português. Em comparação com o PB, o PE revelou uma menor distância acústica do valor F₁ entre /e/ e /ɛ/ - cf. Figura 5. Concluiu-se que a aproximação dos valores de F₁ de /e/ e /ɛ/ deve-se, maioritariamente, à elevação de /ɛ/, isto é, menor abertura da boca, e não tanto à descida de /e/, ficando na posição mais alta no triângulo vocálico. Os autores observaram este fenómeno da convergência iminente de /ɛ/ em /e/, considerando-a uma possibilidade num processo de mudança linguística (Escudero *et al.*, 2009: 1390). Esta observação, relacionada com a convergência das duas vogais médias anteriores, poderá também ser confirmada na percepção dos falantes deste dialeto.

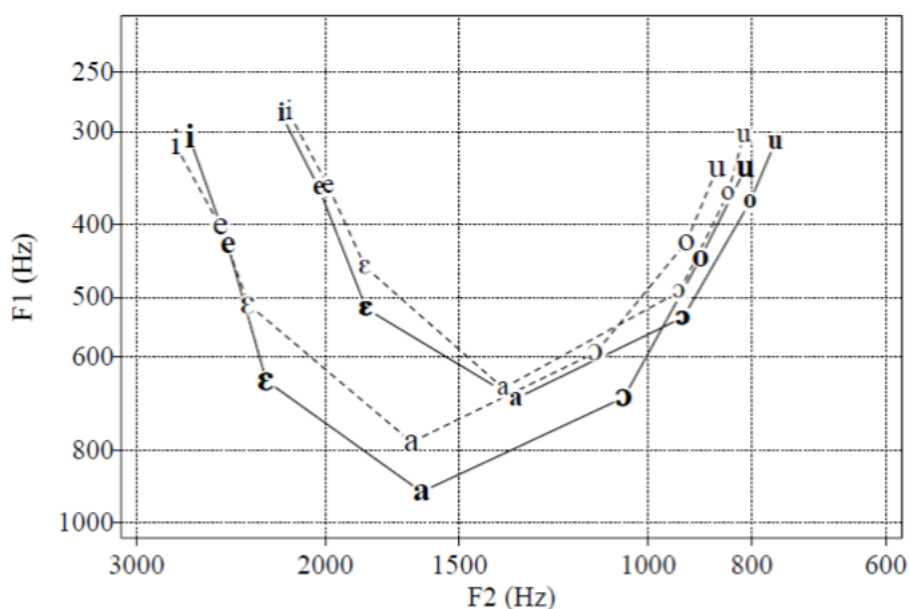


Figura 5: O espaço vocálico de 4 grupos: PB = a linha contínua com símbolo negrito; PE = a linha pontilhada; Homens = as letras maiúsculas; Mulheres = as letras minúsculas. Adaptado de Escudero *et al.* (2009: 1385)

Capítulo 3. - Os estudos experimentais sobre a influência das relações fonológicas na percepção da fala

Neste capítulo iremos abordar os mais recentes estudos alusivos à influência das relações fonológicas na percepção, analisando, com uma visão crítica, as metodologias utilizadas nesses estudos. No fim, refere-se uma série de estudos comparativos entre a fala natural e sintetizada, de forma a sustentar a metodologia escolhida para gerar estímulos no presente estudo.

3.1. As relações fonológicas e a percepção da fala

Recentemente, foram realizados alguns estudos experimentais relativos à percepção da fala com o objetivo de averiguar a influência das relações (com diferentes tipos de complexidade) entre as formas teóricas e as formas de superfície. Hume e Johnson (2003) estudaram o sândi tonal do mandarim, focando-se na relação de contraste parcial, para verificar se existem diferenças perceptivas relativamente à relação de contraste total. O mandarim, uma língua em que o tom tem uma função distintiva, possui quatro tipos de tons lexicais: *level high* (55), *mid-rising* (35), *low-falling-rising* (214) e *high-falling* (51)²⁵. Quando se encontra uma sequência como *low-falling-rising* (214), seguido de *low-falling-rising* (214), o primeiro tom simplifica-se como *mid-rising* (35), neutralizando-se, deste modo, o contraste entre estes dois tons. Por exemplo, /hao²¹⁴ mi²¹⁴/ → [hao³⁵ mi²¹⁴] <bom arroz>. Relativamente a este processo de simplificação, Huang (2001) defendeu que a escolha do tom 35 é perceptualmente tolerada devido à sua maior semelhança com o tom 214, em relação aos outros tons. Nesta experiência participaram falantes nativos de mandarim (em que o tom é fator distintivo) e de inglês americano²⁶ (cujo sistema linguístico não utiliza o tom como fator distintivo). Ambos os grupos realizaram uma tarefa de discriminação AX com os quatro tons produzidos por um falante nativo de mandarim. A distância perceptual foi calculada com o valor do tempo de reação de cada indivíduo. Essa distância perceptual foi posteriormente analisada numa escala

²⁵ Os números que estão entre parênteses indicam os valores de *pitch* dos tons numa escala de cinco níveis (cf. Hume & Johnson, 2001:10)

²⁶ Não se encontra a descrição dos participantes, por exemplo, o número dos elementos de cada grupo, a idade, sexo, etc.

multidimensional para visualizar os espaços perceptuais dos tons de ouvintes chineses e americanos- cf. Figura 6.

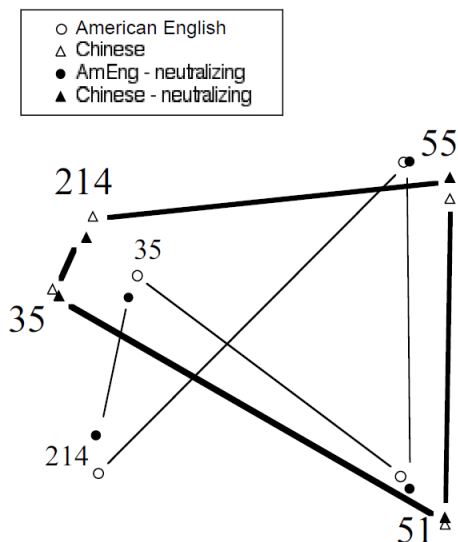


Figura 6: Os quatro tons do mandarim em espaços perceptivos para os ouvintes de chinês mandarim e de inglês americano. Adaptado de Hume & Johnson (2003)

Confirmou-se, através dos resultados de ambos os grupos, a hipótese de Huang (2001) de que o tom 35 é perceptivamente mais semelhante ao tom 214 do que outros tons. Verificou-se também que o espaço perceptivo dos ouvintes chineses é maior do que o dos ouvintes americanos. Esta expansão mostrou que os ouvintes chineses conseguem distinguir, melhor do que os americanos, diferenças tonais na fala. Sendo estes quatro tons lexicalmente distintivos no mandarim, este era um resultado esperado. Salientou-se, no entanto, que a fusão perceptual dos tons 35 e 214 foi observada, não só no contexto neutralizável, mas também no contexto não neutralizável com os ouvintes chineses. Posto isto, concluiu-se que a distinção perceptiva se deve à função dos contrastes fonológicos e que o contraste parcial reduz a sensibilidade dos nativos na distinção entre estes tons, mesmo não se encontrando no contexto neutralizável.

Boomershine *et al.* (2008) estudaram a percepção de sons alofônicos com o objetivo de verificar se existe impacto distinto na percepção do ouvinte, dependendo do tipo de relação existente entre os sons da sua língua nativa, tal como Troubetzkoy (1969 [1939]) tinha já referido. Neste estudo, participaram 10 falantes nativos de espanhol (sexo feminino = 7) na 1ª e 2ª experiências e 7 (sexo feminino = 5) na 3ª e 4ª, todos eles oriundos de países de língua nativa espanhola (nomeadamente, México, Colômbia, Espanha, Porto Rico, Argentina e Peru); e 18 falantes nativos de inglês (sexo feminino = 15) na 1ª e 2ª experiências e 10 (sexo feminino = 7) na 3ª e 4ª. Os falantes nativos de espanhol autoavaliaram a sua competência linguística de inglês segundo uma escala de 1 a 7, sendo que 7 é o nível correspondente ao nativo. Verificou-se, assim, uma classificação média de 5²⁷. Nenhum dos falantes nativos de inglês tinha conhecimento da língua espanhola. Foram estudadas as consoantes [ð, d, r], dado que estas estabelecem relações diferentes nas duas línguas: relação contrastiva a nível fonémico, relação contrastiva a nível fonético e relação alofônica. Em inglês [d, r] são sons alofônicos do mesmo fonema, enquanto que em espanhol este par é contrastivo (embora no nível fonético não se oponham²⁸). O par [d, ð] é sempre contrastivo em inglês (quer no nível de superfície, quer no nível subjacente), enquanto que em espanhol estes são alofones do mesmo fonema. Por último, em ambas as línguas a relação entre [ð] e [r] é contrastiva, contudo, em ambas as línguas, um dos sons é alofone da forma subjacente. Isto é, o contraste fonético [ð] e [r] em inglês corresponde ao contraste fonémico /ð/ e /d/, enquanto o contraste fonético [ð] e [r] em espanhol corresponde ao contraste fonémico /d/ e /r/. Dois foneticistas americanos (1ª e 2ª experiência) e falantes nativos de grego²⁹ (3ª e 4ª experiência) produziram um conjunto de estímulos, no contexto entre as mesmas vogais: [ada], [ara], [aða], [idi], [iri], [iði], [udu], [uru] e [uðu]. Os falantes foram gravados a tentar produzir a primeira e a segunda sílaba com a mesma força acentual, sendo o pico da amplitude controlado de modo a garantir a igualdade de todos os *tokens*. Após a gravação, foram selecionados os dois melhores exemplares para cada sequência VCV³⁰. Foram realizados dois tipos de experiências, pretendendo que os ouvintes operassem processamentos distintos relacionados com cada nível: o nível fonológico e o fonético. Para o processamento fonológico, foi pedido aos ouvintes que classificassem a similaridade perceptiva de um par não idêntico: [d]/[r], [ð]/[d] e [ð]/[r], na escala de 1 a 5. Para o processamento fonético, foi pedido que discriminassem se dois sons apresentados eram iguais ou diferentes (podendo haver um par da mesma sequência com valores físicos diferentes). Nesta última tarefa, os ouvintes foram incentivados a responder rapidamente e com precisão. Os resultados de ambos

os tipos de experiência mostraram um padrão semelhante. Os falantes de espanhol perceberam melhor o contraste entre [d] e [r] - que são fonemas dessa língua, do que entre [ð] e [d]. Enquanto que os falantes de inglês perceberam melhor o contraste entre [ð] e [d] – que são fonemas do inglês, do que entre [d] e [r]. Ou seja, os nativos de cada língua revelaram menor sensibilidade à diferença entre os sons alofônicos do que entre os sons contrastivos da sua língua nativa, nos dois tipos de experiência. No que se refere ao par [ð, r], não se confirmou diferença significativa entre os dois grupos de participantes em qualquer das experiências. Os padrões semelhantes observados em todas as experiências levaram os autores a considerar uma influência maior das relações fonológicas da língua nativa que podem emergir mesmo no nível fonético. Salienta-se, ainda, que se verificou, com os falantes de inglês, entre o mesmo contraste fonémico, uma melhor percepção dos sons que não alteram a sua especificação, do que quando um dos sons é alofone. Os autores propuseram duas explicações para este último resultado: 1) a razão pela qual a melhor percepção do som que não altera a sua especificação reside nas qualidades auditivas gerais dos sons; 2) ou a noção do contraste na percepção é mais subtil. A questão ficou, no entanto, em aberto.

Relativamente ao português, destacam-se dois estudos de Silva e Neves (2009, 2016) sobre o PB, cujo sistema fonológico (tal como o do PE), possui as oposições das vogais médias na sílaba tónica, tendo, assim, sete vogais no seu sistema vocálico. No que concerne ao processo de neutralização das vogais médias, o PB apresenta diferentes realizações em relação às do PE

²⁷ Boomershine (2013) realizou um estudo experimental em que comparou a percepção das vogais inglesas dos 5 grupos com *background* linguístico diferente: monolíngue inglês, monolíngue espanhol, bilingue inglês nativo com espanhol avançado, bilingue espanhol nativo com inglês avançado e falantes de espanhol como língua de herança. Observou-se a influência da experiência linguística na percepção, isto é, quanto mais tempo de exposição exposto a L2, melhor a distinção entre o contraste inexistente no sistema de L1. Por outro lado, no estudo referido neste capítulo (Boomershine *et al.*, 2008), considerando o nível avançado de inglês dos participantes que eram falantes nativos de espanhol, para evitar, tanto quanto possível, a influência da segunda língua, as experiências foram conduzidas em espanhol. Além disso, os estímulos utilizados não foram palavras inglesas. Tendo em conta os resultados obtidos nessa experiência, isto é, as diferenças perceptivas entre os falantes de inglês e os de espanhol, os autores consideraram que os falantes de espanhol processaram cognitivamente em espanhol.

²⁸ Em espanhol, /d/ e /r/ são fonemas. No entanto, não surgem no mesmo contexto. O /d/ aparece em início da palavra, e /r/ no meio (cf. Boomershine *et al.* 2008).

²⁹ Após a 1^a e a 2^a experiências, os investigadores replicaram-nas (3^a e 4^a experiência) com um conjunto de estímulos produzido por falantes gregos, cuja língua possui todas as oposições entre os sons tratados nessa investigação, pretendendo, deste modo, confirmar a mudança na percepção dos participantes (cf. Boomershine *et al.* 2008).

³⁰ Não se encontra uma descrição relativa aos estímulos. Contudo, supõe-se que, através das descrições das experiências 1 e 2, os dois exemplos da mesma sequência são fisicamente diferentes.

no contexto átono não final: [e] para anteriores e [o] para recuadas³¹ (Câmara Jr, 1977; Wetzels, 1992; Bisol, 2001). Silva e Neves (2009) problematizaram o estatuto fonológico do contraste entre as vogais médias dado que estas vogais revelam comportamentos como: 1) inexistência de oposição entre os verbos, podendo somente haver oposição entre os verbos e os nomes cujos radicais possuem o mesmo sentido lexical³²; 2) existência de variações da alternância do grau da abertura na sílaba tónica, no PB, sem mudança do significado, por exemplo, p[o]ça / p[ɔ]ça, cr[o]sta / cr[ɔ]sta, av[e]ssas / av[ɛ]ssas, [e]xtra / [ɛ]xtra; 3) previsibilidade das realizações conforme o contexto, por exemplo, fenómenos como o abaixamento espondeu e o abaixamento datílico (Wetzels, 1992). Tendo em consideração estes comportamentos problemáticos das vogais médias, os autores pretenderam explorar, através do método experimental, possíveis diferenças nos modos como estes contrastes são representados no sistema perceptivo dos falantes de PB. A experiência, de natureza categorial, teve como objetos de investigação as categorias e os limites de categorias do contraste entre vogais médias recuadas [o] e [ɔ]. Quarenta e duas vogais sintetizadas num *continuum* entre [u] e [ɔ] foram geradas com o Praat 4.3.12 (Boersma & Weenink, 2005), variando os valores de F₁ de 10 Hz em 10Hz entre 250-660 Hz. Os valores dos outros formantes (do F₂ ao F₁₀), por sua vez, foram fixados³³. Além disso, o valor da duração (300 ms) e o trajeto da frequência fundamental (f₀) (decaindo linearmente de 145 a 85 Hz), foram determinados de modo a obter sons semelhantes às vogais produzidas pela voz masculina. O contínuo de 42 sons sintetizados foi dividido, posteriormente, em dois *subcontinua* compostos por 21 sons: [u]-[o] 250-470 Hz; [o]-[ɔ] 440-660 Hz. Doze falantes (sexo feminino =6) nativos de PB de Belo Horizonte, com idades compreendidas entre os 18 e os 27 anos, participaram nesta experiência. A experiência incluía uma tarefa de classificação de um *continuum* de sons entre [u] e [ɔ], e duas tarefas de discriminação com dois *subcontinua* entre [u] e [o] por um lado, e entre [o] e [ɔ] por outro: 2AFC (*two alternatives forced choice*) - o sujeito deve indicar a ordem das categorias dos dois sons apresentados, sendo informado previamente das categorias em questão (Silva & Neves, 2009: 324); 4I2FC (*four intervals two*

³¹ Referindo-se à tipologia de neutralização de Troubetzkoy (1967 [1939]: 77), Bisol & Veloso (2016) argumentam que estes processos do PB e PE são distintos. Enquanto as neutralizações das vogais médias anteriores do PE podem ser classificadas como tipo (1b) e das recuadas (3b) -cf. Capítulo 2, secção 2.3. *Propostas teóricas do inventário fonológico das vogais do PE*, as neutralizações das vogais médias do PB pertencem ao tipo (3a), isto é, no contexto neutralizável, a oposição privativa que envolve os traços marcado e não marcado realiza-se como não marcado (Bisol & Veloso, 2016: 70).

³² Veloso (2016: 646) classificou este tipo de oposições como oposições gramaticais.

³³ Os valores fixados: F₂= 900 Hz, as frequências de F₃ (= 2500Hz) a F₁₀ (= 9500Hz) a cada 1000 Hz.

alternatives forced choice) - o sujeito deve identificar, entre dois pares apresentados, qual o par que contém dois estímulos idênticos, sendo informado previamente da possibilidade do estímulo diferente aparecer na segunda ou na terceira posição (Silva & Neves, 2009: 325). Na tarefa de classificação, observou-se uma curva mais abrupta na transição da categoria entre [u] e [o] do que na transição entre [o] e [ɔ]. Esta diferença observada na inclinação das curvas mostrou uma melhor clareza perceptiva na transição da primeira oposição, em comparação com a da segunda. Relativamente às tarefas de discriminação, confirmando-se uma correlação com o resultado da classificação apenas na tarefa de 2AFC³⁴, foi calculado o valor *d'* (*discriminability*) desta tarefa, usado como medida de sensibilidade (Silva & Neves, 2009: 332). Este valor foi mais elevado no *subcontinuum* entre [u] e [o] do que entre [o] e [ɔ]. Concluiu-se, deste modo, que a distinção entre [o] e [ɔ] é menos estável e menos relevante em comparação com a existente entre [u] e [o] ao nível das categorias armazenadas na memória de longo prazo e empregues nos processos perceptivos de um falante de PB.

Silva e Neves (2016) investigaram também a oposição das vogais médias anteriores /e/ e /ɛ/ do PB, teoricamente mais complexa³⁵. Salientam-se duas observações referidas neste estudo: 1) as oposições das vogais médias são menos comuns do que as oposições entre vogais altas e médias, sendo o sistema vocálico com cinco vogais o mais estável e mais observado nas línguas do mundo (Maddieson, 1984); 2) a aquisição mais tardia das oposições vocálicas fonológicas destas vogais pelas crianças do PB (Bonilha, 2004). As hipóteses lançadas pelos autores são as seguintes: 1) a distinção entre /i/ e /e/ será mais nítida do que /e/ e /ɛ/; 2) na categorização, o tempo de resposta perto da fronteira entre /e/ e /ɛ/ será maior do que entre /i/ e /e/ nas mesmas condições. Vinte e oito vogais sintetizadas foram geradas com o sintetizador *KLSYN88* (Klatt & Klatt, 1990), fazendo variar, em intervalos iguais, F₁, F₂ e F₃ ao longo do

³⁴ Na tarefa de 4I2FC verificou-se uma não correlação com o resultado da classificação. Segundo os autores, este facto pode significar que os sujeitos empregaram estratégias de avaliação baseadas nos traços mais sensoriais e não estratégias baseadas nas categorias, uma vez que a correlação entre a identificação e a discriminação é um dos critérios da perceção categorial – cf. Capítulo1, ponto 1.2.2. *Propriedades acústicas e perceção das vogais*.

³⁵ Os autores basearam-se em dois modelos fonológicos de modo a argumentar a complexidade da oposição entre vogais médias: Teoria da Hierarquia Contrastiva (Dresher, 2009) e o modelo de Lee (2010) para PB; Fonologia dos elementos (Harris, 1994; Harris & Lindsey, 2000), analisado por Nevins (2012) em PB. Em ambos os modelos, as descrições destas oposições são mais complexas (cf. Silva & Neves, 2016: 358-359).

continuum de /i/ a /e/³⁶. Estes valores foram baseados nos dados obtidos dos falantes adultos masculinos de PB nos estudos de Escudero *et al.* (2009) e de Rauber (2008). A duração (150 ms) e o trajeto da frequência fundamental (f_0), decaída linearmente de 120 a 90 Hz, foram determinados de modo a que os estímulos fossem semelhantes à voz humana produzida por um falante masculino. Os outros cinco formantes (de F_4 a F_8) foram fixados³⁷. Quarenta e dois falantes (sexo feminino = 20) nativos de PB (falado na região central de Minas Gerais) participaram nesta experiência com duas tarefas de identificação³⁸: 2AFC (escolha forçada entre duas opções) entre /i/ e /e/, por um lado, e entre /e/ e /ɛ/, por outro. O tempo de resposta e a *discriminability* foram analisados estatisticamente. Encontrou-se, nos resultados, uma maior *discriminability* e resposta mais rápida na tarefa de oposição entre /i/ e /e/. Estes resultados indicaram que a fronteira entre /e/ e /ɛ/ é menos definida, portanto, exige um processamento mais complexo na categorização, justificando as descrições fonológicas complexas dessas vogais.

Nos estudos referidos acima, cujos objetivos eram semelhantes ao nosso, obtiveram-se resultados positivos no que toca à confirmação da seguinte hipótese: as relações estabelecidas entre fonemas no sistema fonológico influenciam a percepção. No entanto, é pertinente fazer algumas observações acerca destes estudos. Tendo em consideração a especificidade da língua e a plasticidade da percepção - cf. ponto 1.1.3 *Especificidade da língua*, em Boomershine *et al.* (2008), a falta de homogeneidade dos participantes poderá ter influenciado os resultados do estudo. O mesmo pode ser dito do facto de não terem sido usados, em Silva e Neves (2016), estímulos de voz natural³⁹. Além disso, a teoria utilizada para desenvolver o nosso estudo não corresponde à usada por Silva e Neves (2016). Tal como mencionado na introdução, a teoria utilizada no nosso estudo é a teoria do protótipo - cf. secção 1.3. *Teoria do protótipo: estrutura*

³⁶ “The three lower formant frequencies varied along the vowel *continuum* in fixed steps of 0.13 (F_1), 0.06 (F_2) and 0.03 (F_3) Bark, from 2.01 to 5.39 (F_1), 14.20 to 12.52 (F_2), and 15.69 to 14.76 Bark (F_3) – or, in Hertz, from 205 to 555, 2390 to 1860, and 3000 to 2600 Hz.” (Silva & Neves, 2016: 361)

³⁷ $F_4=3900$, $F_5=4900$, $F_6=5900$, $F_7=6900$ e $F_8=7900$ Hz

³⁸ Os autores escolheram o termo “classificação” para a tarefa de 2AFC. De modo a evitar confusão com o termo usado na nossa investigação, optamos por usar o termo “identificação” para tal tarefa.

³⁹ Os próprios autores referiram que era interessante realizar o mesmo tipo de investigação com os estímulos gerados a partir da fala natural, lançando a hipótese de que, neste caso, os índices de percepção categorial serão maiores na fala natural. (Silva & Neves 2009: 339-340) O que não coincide com a nossa visão sobre as características das vogais, isto é, quanto mais informação fonética for transmitida, os ouvintes poderão ficar mais sensíveis às diferenças dentro da categoria. Assim sendo, com estímulos gerados a partir da fala natural, demonstrar-se-ia um menor grau do paradigma da percepção categorial.

interna da categoria dos sons da fala, de forma a fornecer evidências sob um novo ponto de vista.

Na seguinte secção, referimos uma série das experiências da *Haskins Laboratories* em que compararam a voz sintetizada e a natural, de forma a argumentar a metodologia escolhida para gerar estímulos no presente estudo.

3.2. Criação dos estímulos nas experiências da percepção de fala

Tendo em consideração a melhor inteligibilidade da fala natural face à voz sintetizada, uma das questões metodológicas a ter em conta antes de desenhar qualquer experiência de percepção está ligada à natureza/características dos estímulos. Estas, pois, podem influenciar o resultado numa experiência que trate da percepção de fala. Há, efetivamente, estudos experimentais anteriores que mostram uma diferença perceptual entre a fala natural e a sintetizada, tanto na sua inteligibilidade como na sua compreensão. Em Logan, Greene & Pisoni (1989), observou-se, no teste de rima modificada com as palavras monossilábicas CVC do inglês, a melhor inteligibilidade dos segmentos na fala natural (melhor que 99%) em comparação com a fala gerada por 10 sintetizadores⁴⁰ diferentes (variando de 63% a 96%). A maior inteligibilidade da fala natural comparada com a fala sintetizada foi também confirmada na segunda língua. Em Greene (1986), os participantes não nativos do inglês tiveram melhor desempenho com a fala natural nas respostas relativas aos segmentos, palavras, frases e texto. Em Pisoni *et al.* (1987) realizou-se uma experiência onde se verificou a compreensão das frases produzidas com fala natural e sintetizada. Nessa experiência, os ouvintes tinham que identificar se a frase ouvida era verdadeira ou falsa. Por exemplo: <Cotton is soft> ou <Snakes can sing>. Os participantes responderam consistentemente mais rápido com fala natural. Assim, os resultados desta experiência mostraram melhor compreensão da fala natural em comparação com a fala sintetizada. Pisoni (1997: 546) concluiu que, mesmo que a voz sintetizada tenha alta qualidade e inteligibilidade, os dois tipos de voz não são processados cognitivamente da mesma maneira, isto é, a voz sintetizada requer um esforço adicional para a compreensão. Este mesmo fenómeno foi também confirmado na tarefa de *word-monitoring* de Ralston *et al.* (1991). Os participantes conseguiram detetar mais rápido as palavras-alvo memorizadas anteriormente

⁴⁰ Usaram-se os seguintes sintetizadores: MITalk-79, TSI-Prototype, Infovox SA 101, Prose 2000 V3.0, Votrax Type 'n'Talk, Echo. Amiga, Smooth talker, DECtalk 1.8 Paul, e DECtalk 1.8 Betty.

num excerto produzido com fala natural. Pisoni (1997: 551) inferiu que esta dificuldade no processamento da fala sintetizada poderá ter origem na falta de redundância adicional e variabilidade acústico-fonética encontrada na fala natural, ou seja, enquanto a fala natural contém informações acústicas ricas e abundantes para cada contraste fonético, a sintetizada possui apenas um número mínimo destas pistas acústicas⁴¹. Embora ainda não estejam bem esclarecidos os papéis da redundância adicional e das variabilidades acústicas na percepção da fala, a sua complexidade deverá ser um fator importante.

Tendo em consideração a afirmação de Pisoni (1997), e também os resultados dos estudos anteriores acima referidos, escolhemos, no nosso estudo, gerar estímulos a partir da voz humana com o *software TANDEM-STRAIGHT* (Kawahara *et al*, 2009), que nos permite gerar voz sintetizada com as informações abundantes da fala natural. Abordamos, seguidamente, a função desse *software*.

3.3. TANDEM-STRAIGHT

O *software TANDEM-STRAIGHT* (Kawahara *et al*, 2009) foi desenvolvido por Kawahara e colegas, incorporando os mais recentes resultados do modelo de audição e processamento do sinal de fala, de modo a realizar uma voz sintetizada com alta qualidade dificilmente distinguível da fala natural. O desenvolvimento do programa *STRAIGHT*, modelo precedente do *TANDEM-STRAIGHT*, começou em 1997 com o objetivo de fornecer uma ferramenta para experiências da percepção de fala. A sua estrutura baseia-se no Modelo Fonte Filtro da produção de fala⁴² (Fant, 1960), ou seja, segundo estes investigadores, “*STRAIGHT is essentially a contemporary version of a channel VOCODER*” (Kawahara *et al.*, 2009:111) e “*a program that analyzes speech, decomposes it into source and filter characteristics, and then*

⁴¹ Salienta-se que, em Schwab *et al.* (1985), verificou-se o melhoramento do desempenho nas 5 tarefas após treino com fala sintetizada. Verificou-se que os ouvintes são capazes de modificar as suas estratégias perceptivas, num tempo relativamente curto, ao se familiarizarem com fala sintética de baixa qualidade em comparação com a fala natural.

⁴² A teoria fonte-filtro (teoria acústica da produção da fala) foi desenvolvida inicialmente por Fant (1960) e posteriormente continuada por outros investigadores. De acordo com esta teoria, as características acústicas dos sons da fala são definidas pelas articulações do aparelho fonador. As fontes sonoras que geram som podem ser periódicas (resultado da vibração das cordas vocais) ou aperiódicas (resultado do ruído gerado ao longo do trato vocal). As características dos filtros, que modelam o trato vocal, são determinadas pelas configurações das cavidades supraglóticas que modificam a estrutura da onda sonora (Falé, 2005: 120).

recombines it after some manipulation or interpolation.” (McAuliffe, 2016). A decomposição praticamente completa do sinal de fala (*input*) em informações da fonte e as espectrais (que são as informações de filtro) possibilita uma manipulação flexível dos parâmetros da fala com uma degradação mínima da qualidade da fala sintetizada, que era o maior problema dos sintetizadores anteriores baseados no modelo fonte-filtro (Kawahara *et al.*, 2009:111). Deste modo, este *software* permite produzir estímulos sintetizados da forma mais natural possível, isto é, são mantidas as informações ricas da fala natural. Na Figura 7, pode observar-se a fala gerada com um sintetizador normal de formantes, cujos espectros não possuem picos extra perto de formantes principais. O *STRAIGHT*, por outro lado, produz um som sintetizado com variabilidade acústica, comprovado no espectro pelos vários picos extra existentes (Liu & Kewley-Port, 2004: 1231), que é uma das características fundamentais da fala natural.

Uma das funções essenciais do *TANDEM-STRAIGHT* é a de fazer *morphing* entre dois sons, de modo a gerar estímulos contínuos. O termo “*morphing*” foi inicialmente usado no contexto da tecnologia da computação gráfica para descrever a transformação de uma imagem noutra, através de uma transição de modo gradual e natural. Na criação dos sons sintetizados, portanto, *morphing* significa produzir sons intermédios entre duas amostras de forma contínua e natural, tal como na transição da imagem dos efeitos especiais nos filmes.

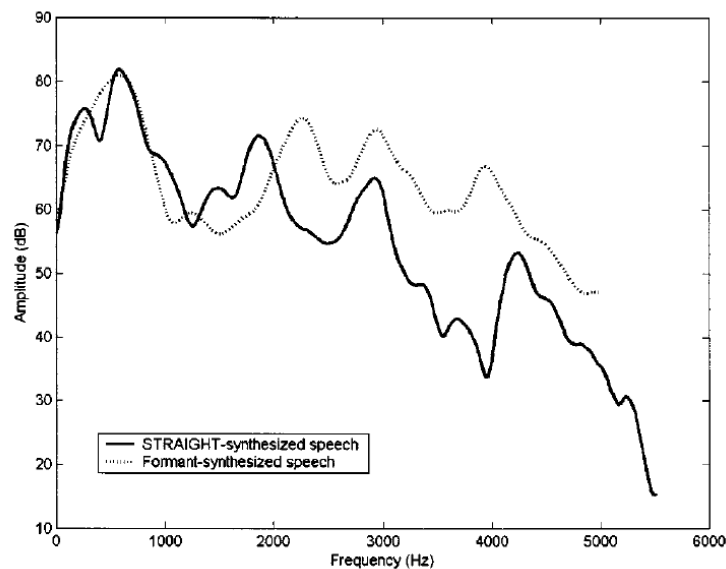


Figura 7: Espectros da vogal /ε/ gerados pelo *STRAIGHT* (Kawahara *et al.* 1999) e pelo sintetizador Klatt (1980). Destaca-se o espectrograma gerado com o *STRAIGHT* que não representa uma boa estrutura harmónica nem a propriedade da fonte, que são armazenadas noutros parâmetros do programa. Adaptado de Liu & Kewley-Port (2004: 1231).

O procedimento *morphing*, usado pelo *TANDEM-STRAIGHT*, desenha uma trajetória entre duas amostras através do cálculo individual dos dados extraídos de três informações (traçado de f_0 , espectrograma e espectrograma *STRAIGHT*). Para tal, o *TANDEM-STRAIGHT* realiza: o alinhamento de duas amostras; a configuração da taxa de *morphing* e a criação do *continuum*. Na primeira etapa, confirmando visualmente a distância (calculada pelo programa), entre duas amostras, alinham-se, temporalmente, as duas amostras da forma de onda e também do domínio da frequência. Em seguida, marcam-se os pontos inicial e final e introduz-se a taxa de *morphing*, onde se podem configurar vários parâmetros, dependendo do que se pretende. A partir do momento em que se marcam estes pontos, pode criar-se um *continuum*. O *software* gera o contínuo com o número de estímulos que se pretender, sendo que, por defeito, o programa escolhe 11.

Capítulo 4. - Metodologia

Neste capítulo é descrita a metodologia adotada neste estudo, incluindo questões relacionadas com a criação de estímulos, a seleção das técnicas de recolha de dados e a descrição e avaliação dos instrumentos de recolha de dados.

4.1. Estímulos

Tal como referido na secção 3.2. *Criação dos estímulos nas experiências da percepção de fala*, os estudos anteriores revelaram uma diferença perceptiva entre a fala natural e a fala sintetizada, devido à falta de informações acústicas (típicas da fala natural) na fala sintetizada (Greene, 1986; Logan, Greene & Pisoni, 1989; Ralston *et al.*, 1991; Pisoni, 1997). Além disso, existe uma forte hipótese da existência de uma correlação entre a produção e a percepção (Newman, 2003; Evans & Iverson, 2004, 2007). Hipótese essa que se mostra pertinente, tendo em conta a observação e a questão colocada no estudo de Escudero *et al.* (2009), isto é, a convergência das vogais médias anteriores na produção dos jovens lisboetas. Essa convergência pode ser também verificada na percepção. Considerando estes fatores, de modo a obter dados fiáveis, os nossos estímulos foram gerados de acordo com os seguintes critérios: 1) selecionar um informante masculino⁴³ que represente a fala típica de Lisboa; 2) manter as características da fala natural durante o processo de síntese. Assim sendo, realizaram-se as quatro seguintes etapas: 1) gravação de amostras de fala de jovens lisboetas; 2) inquérito online para selecionar um dos informantes; 3) gravação das vogais-alvo; e 4) criação dos estímulos.

4.1.1. Seleção da voz de Lisboa

Postulando correlações entre a produção e a percepção, fizemos a seleção de um informante com as mesmas características do estudo de Escudero *et al.* (2009) isto é, um jovem

⁴³ Em Escudero *et al.* (2009), a convergência das vogais médias foi confirmada em ambos os sexos. Pretendendo fazer uma comparação entre os resultados de Silva e Neves (2016), em que investigaram o mesmo tema com o PB, e os nossos, optámos por utilizar a voz masculina para esta investigação.

masculino com residência permanente em Lisboa⁴⁴, de modo a gerar os estímulos a partir da fala natural. Os cinco informantes que satisfizeram os critérios mencionados acima participaram na gravação (Tabela 4). Assinaram também uma declaração, que mostra a sua participação voluntária e autorização à gravação da sua fala natural para fins académicos, sobre a condição de anonimato (cf. Anexo 2).

Tabela 4: Dados sociodemográficos dos falantes.

Falante	Sexo	Idade	Naturalidade	L1	Estatuto social	Grau académico	Observações
1	M	21	Lisboa	PE	Estudante universitário	3º ano da Licenciatura	
2	M	28	Lisboa	PE	Estudante universitário	2º ano da Licenciatura	
3	M	28	Lisboa	PE	Trabalhador	Mestrado	
4	M	25	Lisboa	PE	Trabalhador	Mestrado	Mãe é japonesa
5	M	35	Lisboa	PE	Investigador	Doutoramento	Esteve 6 meses no Porto

As amostras de fala foram gravadas, individualmente, em sítios silenciosos de Lisboa e do Porto, com um microfone *TRUST Mico USB* ligado a um computador portátil. As gravações foram realizadas com o programa *Praat (version 6.0.37)*, a uma frequência de amostragem de 48000 Hz, com 16 bits por amostra, e os dados armazenados em formato mono .wav (Windows PCM) sem compressão. Foi pedido aos participantes para falarem espontaneamente durante 2 minutos, sobre um tema comum (futebol), esperando, deste modo, uma amostra de fala o mais natural possível.

4.1.2. Inquérito online

As cinco gravações (com a duração total), foram colocadas separadamente no sítio *web SoundCloud*, criando de seguida ligações a um formulário do Google, intitulado “As vogais do português de Lisboa” – cf. Figura 8.

⁴⁴ Nesta fase, deixamos de parte critérios como o *background* familiar e o nível de proficiência noutras línguas, que ficam para análise pelos peritos.

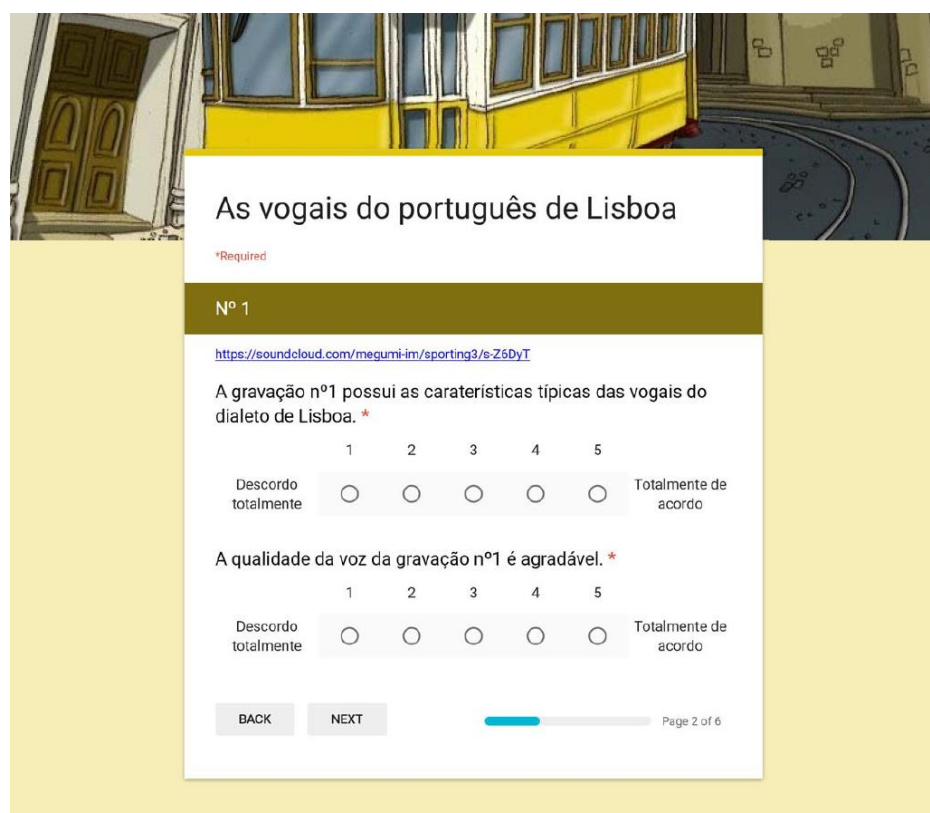
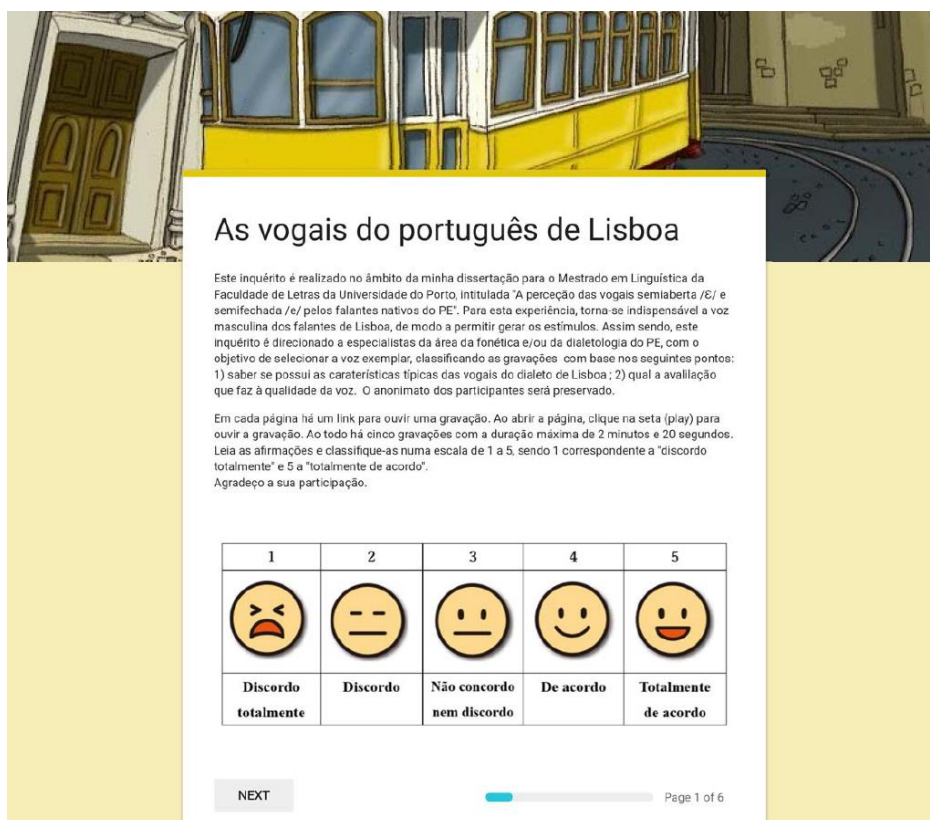


Figura 8: Captura de ecrã das páginas de avaliação do inquérito online.

Realizamos, então, um inquérito online dirigido a especialistas da área da fonética e/ou da dialetologia do PE, com o objetivo de selecionar uma amostra de fala exemplar de Lisboa. As gravações foram classificadas por 8 especialistas, através de uma página *web*, numa escala de 1 a 5, sendo 1 correspondente a “discordo totalmente” e 5 a “totalmente de acordo”. As afirmações do inquérito são as seguintes: 1) a gravação nº X possui as características típicas das vogais do dialeto de Lisboa; 2) a qualidade da voz da gravação nº X é agradável. Conforme se constata na Tabela 5, as gravações nº 1 e nº 4 obtiveram a melhor classificação. No que toca às características típicas das vogais do dialeto de Lisboa, essas duas gravações obtiveram a mesma pontuação. Em termos da qualidade vocal, houve uma pequena diferença entre essas. Tendo em conta a natureza do nosso estudo, em que foi pedido aos participantes para ouvirem os estímulos repetidamente durante 45 minutos (no mínimo), optamos por selecionar a voz considerada mais agradável. O falante 1 foi selecionado com a classificação da média mais alta associada à menor variabilidade (o menor desvio padrão), considerando também a soma de pontos mais alta.

Tabela 5: Avaliação das gravações pelos 8 peritos da fonética e da dialetologia do PE. Os números na primeira linha representam os números das gravações, apresentados no inquérito online, que correspondem aos números dos falantes da tabela I. a1 representa a afirmação “possui as características típicas das vogais do dialeto de Lisboa” e a2 representa a afirmação “a qualidade da voz da gravação é agradável”

Avaliadores	Nº1		Nº2		Nº3		Nº4		Nº5	
	a1	a2	a1	a2	a1	a2	a1	a2	a1	a2
A	4	4	3	5	3	3	5	4	3	4
B	5	4	4	3	4	4	4	4	5	4
C	5	5	4	4	5	4	5	5	4	4
D	4	4	4	5	4	4	4	4	3	4
E	5	4	4	4	4	5	4	5	4	5
F	4	4	4	3	3	2	5	2	4	2
G	3	5	1	4	3	4	3	4	3	3
H	4	5	3	2	3	5	4	4	3	3
Total	34	35	27	30	29	31	34	32	29	29
a1+a2	69		57		60		66		58	
Média	4.25	4.38	3.38	3.75	3.63	3.88	4.25	4	3.63	3.63
D.P.	0.66	0.48	0.99	0.97	0.7	0.93	0.66	0.87	0.7	0.86
Mediana	4	4	4	4	3.5	4	4	4	3.5	4

4.1.3. Gravação das vogais-alvo

Tendo em conta o resultado obtido na avaliação das gravações pelos 8 peritos da fonética e da dialetologia do PE, realizamos outras gravações em Lisboa com o falante selecionado na etapa anterior. Utilizamos um microfone de condensador *AKG Perception 120 USB* ligado a um computador portátil. As gravações foram realizadas com o programa *Praat (version 6.0.37)*, a uma frequência de amostragem de 48000 Hz, com 16 bits por amostra, e os dados armazenados em formato mono .wav (Windows PCM) sem compressão. Para se habituar e facilitar a pronúncia, foi pedido ao falante para pronunciar uma vez palavras dissilábicas que contêm a vogal em estudo na sílaba tónica, sendo estas: <pico>, <medo>, <teto> e <pato>. Depois de pronunciar a palavra dissilábica, pronunciava três vezes a vogal-alvo em questão isoladamente, sendo estas: [i, e, ε, a]. Optamos por gravar amostras com uma descida do tom no final, um procedimento semelhante ao utilizado no estudo de Silva & Neves (2016), procurando gerar os estímulos da forma mais natural possível. Decidimos, também, utilizar uma duração de 400 ms, semelhante ao estudo de Masapollo *et.al* (2017).

4.1.4. Criação dos estímulos

Selecionamos as quatro vogais, pronunciadas isoladamente, com os seguintes critérios: 1) descida de f_0 no final; 2) valores de F_1 e F_2 o mais próximo possível dos de Escudero *et al.* (2009) (Tabela 6). Os valores das frequências dos formantes foram extraídos no ponto médio, utilizando a seguinte função do *Praat 6.0.37* (valores dos parâmetros escolhidos por defeito no SoundEditor): To Formant (burg)... 0.01 5 5500; 0.025 50; [Time step(s), Max. number of formants, Maximum formant (Hz), Window length(s), Preemphasis from (Hz)] – split Levinson algorithm. Foi também utilizado o espectrograma (valores dos parâmetros escolhidos por defeito no SoundEditor) para verificar os valores extraídos e efetuar algumas correções manuais pontuais. Pode confirmar-se a descida da frequência fundamental (f_0), ± 20 - 25 Hz, do início ao fim, na Figura 9.

Tabela 6: Valores médios de f_0 , F_1 , F_2 e F_3 (Hz) e duração (ms) das vogais [i],[e],[ɛ] e [a] produzidas pelos falantes masculinos de Lisboa no estudo de Escudero et al. (2009: 1383)

	f_0 (Hz)	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)	Duração (ms)
Escudero et al. (2009) [i]	126	284	2161	3283	84
Escudero et al. (2009) [e]	122	355	1987	3007	97
Escudero et al. (2009) [ɛ]	117	455	1836	2943	106
Escudero et al. (2009) [a]	115	661	1365	2535	108

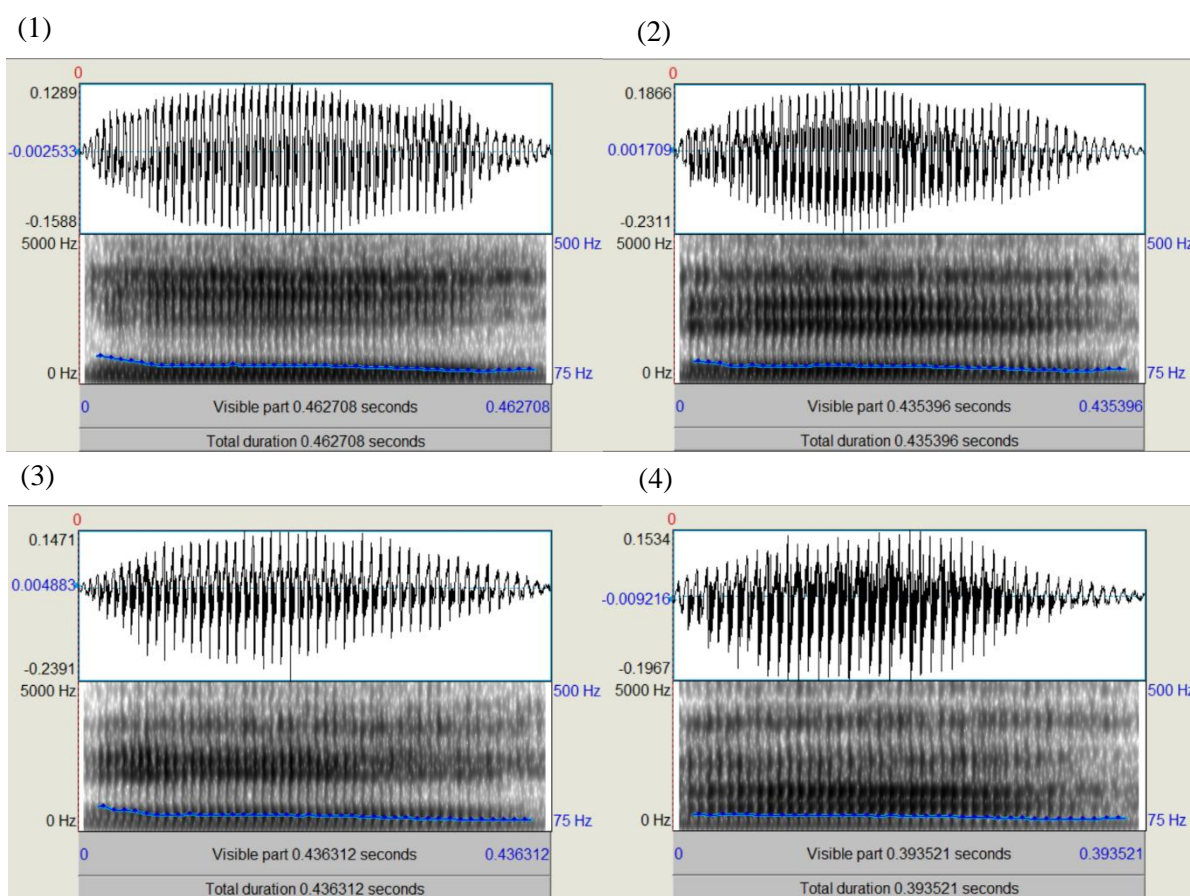


Figura 9: Formas de onda (em cima) e espectrogramas (em baixo) das quatro amostras das vogais orais do português europeu (de Lisboa), produzidas por um falante masculino, medidos com o Praat (version 6.0.37): (1) amostra [i] com a duração de 463 ms, f_0 123 Hz com descida no fim, F_1 319 Hz, F_2 2245 Hz, F_3 2891 Hz. (2) amostra [e] com a duração de 440 ms, f_0 120 Hz com descida no fim, F_1 376 Hz, F_2 1918 Hz, F_3 2599 Hz. (3) amostra [ɛ] com a duração de 440 ms, f_0 117 Hz com descida no fim, F_1 440 Hz, F_2 1860 Hz, F_3 2391 Hz. (4) amostra [a] com a duração de 440 ms, f_0 114 Hz com descida no fim, F_1 709 Hz, F_2 1261 Hz, F_3 2253 Hz.

Estas quatro amostras funcionaram como âncoras, tendo sido gerados, a partir destas, 31 estímulos com o procedimento *morphing* no *TANDEM-STRAIGHT monolithic Package 014* (função *MorphingMenu Last Modified by GUIDE v2.5 19-Jul-2016 01:42:59*), dividindo cada trajetória em 11 etapas. Na Tabela 7 pode verificar-se que, à medida que o valor de F_1 aumenta, os valores de F_2 e F_3 diminuem, refletindo as características do filtro dos estímulos⁴⁵.

Tabela 7: Valores de f_0 , F_1 , F_2 e F_3 (Hz) e duração (ms) dos 31 estímulos sintetizados, obtidos com a parametrização por defeito do Praat (version 6.0.37) SoundEditor.

Nº de estímulos	f0 (Hz)	F1 (Hz)	F2 (Hz)	F3 (Hz)	Duração (ms)
stimulus001	123	319	2245	2891	465
stimulus002	123	319	2192	2860	462
stimulus003	122	323	2167	2854	460
stimulus004	122	329	2134	2788	457
stimulus005	121	335	2074	2775	454
stimulus006	122	344	2049	2754	452
stimulus007	121	352	2030	2720	449
stimulus008	121	360	2001	2699	447
stimulus009	121	366	1972	2668	444
stimulus010	120	371	1936	2609	442
stimulus011	120	376	1918	2599	440
stimulus012	120	381	1906	2580	440
stimulus013	119	389	1899	2547	440
stimulus014	119	397	1895	2533	440
stimulus015	119	406	1893	2503	440
stimulus016	119	415	1891	2491	440
stimulus017	118	423	1885	2462	440
stimulus018	118	430	1872	2452	440
stimulus019	118	434	1867	2420	440
stimulus020	117	440	1864	2393	440
stimulus021	117	440	1860	2391	440
stimulus022	117	468	1842	2348	435
stimulus023	117	492	1771	2343	430

⁴⁵ 15 % (14/93) dos valores dos formantes, obtidos automaticamente, foram manualmente corrigidos, com base no espectrograma e no espectro LPC.

Nº de estímulos	f0 (Hz)	F1 (Hz)	F2 (Hz)	F3 (Hz)	Duração (ms)
stimulus024	116	524	1720	2339	425
stimulus025	116	551	1648	2338	421
stimulus026	115	578	1586	2337	416
stimulus027	116	611	1496	2297	412
stimulus028	115	637	1468	2290	407
stimulus029	115	669	1373	2288	403
stimulus030	115	693	1315	2280	398
stimulus031	114	709	1261	2253	394

4.2. Questionário

Foi realizado um questionário relativo às informações biográficas e perfil linguístico - cf. Anexo 1, de forma a obter um grupo relativamente homogéneo, e comparável com um dos grupos do estudo de Escudero *et al.* (2009). Assim sendo, os participantes foram selecionados com os seguintes pré-requisitos: 1) PE como língua materna; 2) pais portugueses; 3) residência sempre na cidade de Lisboa; 4) inexistência de dificuldades/problemas auditivos; 5) jovem adulto; 6) grau de habilitações académicas acima de licenciado ou frequência numa licenciatura (ensino universitário ou politécnico); 7) inexistência de treino fonético. Relativamente ao perfil linguístico, os participantes autoavaliaram⁴⁶ a sua proficiência em língua(s) estrangeira(s), de acordo com o QECRL -Quadro Europeu Comum de Referência para as Línguas.

4.3. Autorização

Todos os participantes do pré-teste e do teste principal assinaram uma declaração de modo a exprimir a sua participação voluntária e a sua autorização para a recolha e análise de dados, para fins académicos, sob a condição de anonimato – cf. Anexo 2.

⁴⁶ Inglês C2=1, C1=3, B1=1, A2=1, Japonês A2=2, A1=1 no pré-teste e, Inglês C2=5, C1=3, B2=1, A2=1, Inglês C1 e Espanhol B1=1, Inglês B1 e Espanhol B1=1, Inglês C1 e Francês A1=1 no teste principal (nível=nº de participantes).

4.4. Pré-teste

O pré-teste consistiu na aplicação do teste produzido a alguns falantes, que serviu para testar a eficácia do teste e verificar se existia algum problema nos estímulos gerados ou no procedimento da experiência. Neste pré-teste, o ponto 3 do questionário - “se reside sempre na cidade de Lisboa” não foi considerado e apenas se questionou qual a origem dos participantes. Obtivemos um grupo homogéneo, sendo todos da região do Norte, nomeadamente, Braga e Porto. Todas as experiências do pré-teste se realizaram numa sala silenciosa da Universidade do Minho.

4.4.1. Participantes do pré-teste

Seis falantes (sexo feminino =3) nativos de PE, oriundos do Norte de Portugal, estudantes universitários, com idades compreendidas entre os 20 e os 24 anos (média =22.3 anos), que satisfizeram os critérios de inclusão supramencionados, participaram voluntariamente no pré-teste.

4.5. Teste Principal

4.5.1. Participantes

Quinze falantes (sexo feminino =9) nativos de PE, oriundos de Lisboa, com habilitações académicas acima de licenciatura ou a frequentar uma licenciatura, com idades compreendidas entre 19 e 34 anos (média = 23.7 anos) que satisfizeram os critérios de inclusão supramencionados na secção 4.2. *Questionário*, participaram voluntariamente no teste principal.

4.6. Procedimento e *design* da experiência

Os participantes ouviram cada estímulo 20 vezes, ordenados aleatoriamente (620 tentativas na totalidade), com uma pausa opcional após cada 15 tentativas. Não houve possibilidade de repetição da audição. A experiência inclui duas tarefas: 1) escolher uma das palavras apresentadas no ecrã que contém a vogal tónica correspondente ao som ouvido, existindo uma opção X caso não existisse correspondência (tarefa de identificação); 2)

classificar a qualidade do som ouvido numa escala de 1 a 5 conforme a vogal escolhida na tarefa, sendo 1 um mau exemplo desta vogal e 5 um bom exemplo (tarefa de classificação de *goodness*). Para que os participantes compreendessem o procedimento, realizou-se um teste de ensaio com 21 estímulos contínuos de vogais diferentes das visadas pelo estudo, neste caso /u/-/o/, com uma ordem aleatória e sem repetição. Nesse ensaio ajustou-se o volume dos sons individualmente.

As experiências principais realizaram-se em salas silenciosas, na cidade de Lisboa. Os estímulos sintetizados com o *TANDEM-STRAIGHT* (Kawahar *et al.*, 2009) foram apresentados via *Praat* (version 6.0.37), instalado num computador *TOSHIBA dynabook PT65DGP-RJA*, utilizando auscultadores *Sennheiser HD 380 Pro* ligados à placa de som interna do mesmo. Durante a apresentação dos estímulos, as palavras dissilábicas com a vogal-alvo presente na sílaba tónica, <pico>, <medo>, <teto>, <pato> e <X> foram exibidas no ecrã do computador. Além das palavras referidas também foi exposta a escala de 1 a 5 na parte inferior do ecrã – cf. Figura 10.

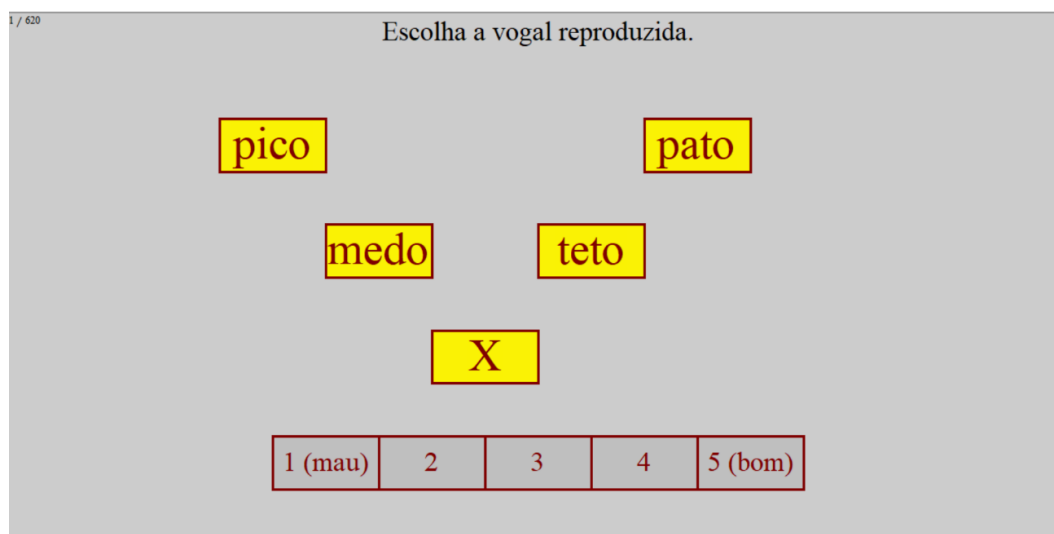


Figura 10: Captura de ecrã da tarefa de identificação e de classificação de *goodness*

As respostas foram dadas usando o rato do computador, seleccionando uma palavra e um número para cada som ouvido. Os dados da identificação (tarefa 1) e da classificação de *goodness* (tarefa 2) foram recolhidos com o *Praat* (version 6.0.37).

4.7. Tratamento dos dados

Os dados foram tratados quantitativamente, realizando uma análise estatística com o *software R studio (Version 1.1.453)* e o *Excel 2016*, sendo os mesmos usados para a criação de gráficos. Visto que se pretende analisar estatisticamente as fronteiras entre categorias na tarefa de identificação (tarefa 1), e as estruturas internas de cada categoria e o grau de nitidez de cada fronteira na tarefa de classificação de *goodness* (tarefa 2), consideramos o número dos estímulos como variáveis independentes, sendo cada um uma identidade distinta definida pelos valores dos parâmetros acústicos, e as respostas dos participantes em cada tarefa como variáveis dependentes.

Capítulo 5. - Resultados

O objetivo deste capítulo é a apresentação e análise dos resultados recolhidos ao longo do estudo efetuado. Inicia-se com a apresentação dos resultados obtidos na tarefa de identificação e seguidamente, dos dados da tarefa de classificação de *goodness*, aplicando a estes dados uma análise estatística efetuada de modo a responder às questões específicas do nosso estudo, mencionadas na *Introdução*.

5.1. Tarefa de identificação

A primeira questão de investigação foi colocada no sentido de se encontrar, na perceção, fronteiras entre as categorias fonémicas /i/-/e/, /e/-/ε/ e /ε/-/a/. Para determinar as fronteiras entre duas categorias, foram utilizados os dados da tarefa da identificação em que os participantes categorizaram 31 estímulos repetidos 20 vezes, por uma ordem aleatória (620 tentativas na totalidade), em 4 categorias vocálicas, podendo também escolher uma opção X caso decidissem que o som ouvido não pertencia a qualquer categoria. A partir dos dados obtidos nesta tarefa, optou-se, primeiro, por uma análise de regressão logística, utilizada também no estudo de Silva & Neves (2016)⁴⁷. O modelo da regressão logística foi escolhido, pois, em primeiro lugar, as respostas da tarefa de identificação não são variáveis contínuas, mas sim variáveis categóricas. Segundo, tratando os dados da identificação como dados binários (isto é, existem apenas dois valores possíveis para a resposta – 0 se pertencer a uma categoria e 1 se não pertencer) pode-se representar, com este modelo, a probabilidade da transição de uma categoria para outra com alta robustez. Este tratamento binário é sustentável, baseando-se na teoria fonética-acústica de que, numa dada zona de dispersão acústica, não é expectável observar respostas de outras zonas não contíguas. E ainda, terceiro, este modelo facilita a interpretação dos resultados, em comparação com o modelo de regressão multinomial, que implica uma interpretação complexa.

Os dados da identificação de cada participante foram representados usando curvas de regressão logística. Estas curvas permitem estimar a fronteira entre duas categorias, que corresponde à probabilidade de resposta $p = 0.5$. Este ponto da fronteira pode ser encontrado usando a equação da regressão logística, quando p toma o valor 0.5:

⁴⁷ As tarefas utilizadas em Silva e Neves (2016) foram a 2AFC (escolha forçada entre duas opções) entre /i/ e /e/, por um lado, e entre /e/ e /ε/, por outro – cf. Capítulo 4 *Metodologia*.

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

onde o parâmetro p representa a probabilidade da resposta não se inserir numa determinada categoria⁴⁸, x é a variável independente que corresponde à identidade representada pelo número de estímulos, e β_0 e β_1 são a ordenada na origem e o declive, respetivamente (cf. Silva & Neves, 2016: 363). Deste modo, o valor de x que corresponde à fronteira entre categorias calcula-se substituindo p por 0.5 de acordo com:

$$\ln\left(\frac{0,5}{1-0,5}\right) = \beta_0 + \beta_1 x_{fronteira}$$

$$\Leftrightarrow 0 = \beta_0 + \beta_1 x_{fronteira}$$

$$\Leftrightarrow x_{fronteira} = -\frac{\beta_0}{\beta_1}$$

$$\text{Fronteira} = -\frac{\beta_0}{\beta_1}$$

Finalmente, para cada participante são estimadas três fronteiras entre as categorias. De seguida podem-se visualizar as Figuras 11, 12 e 13 que correspondem à regressão logística de todos os participantes para a transição das categorias /i/, /e/, /ε/ e /a/ respetivamente.

⁴⁸ Na nossa tarefa de identificação, houve cinco opções de escolha, incluindo “X” caso não existisse correspondência com o som que foi ouvido. Deste modo, p não representa a probabilidade de uma das respostas alternativas, tal como empregue em Silva e Neves (2016), mas a probabilidade de não ser uma determinada resposta.

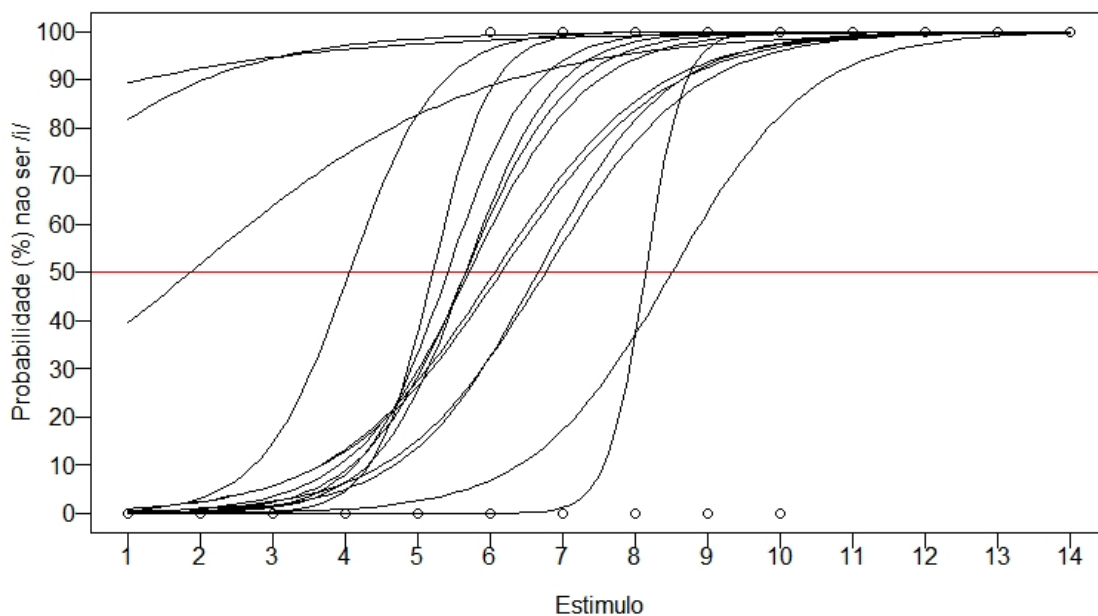


Figura 11: As curvas de regressão logística de todos os participantes para a transição da categoria /i/ para não /i/.

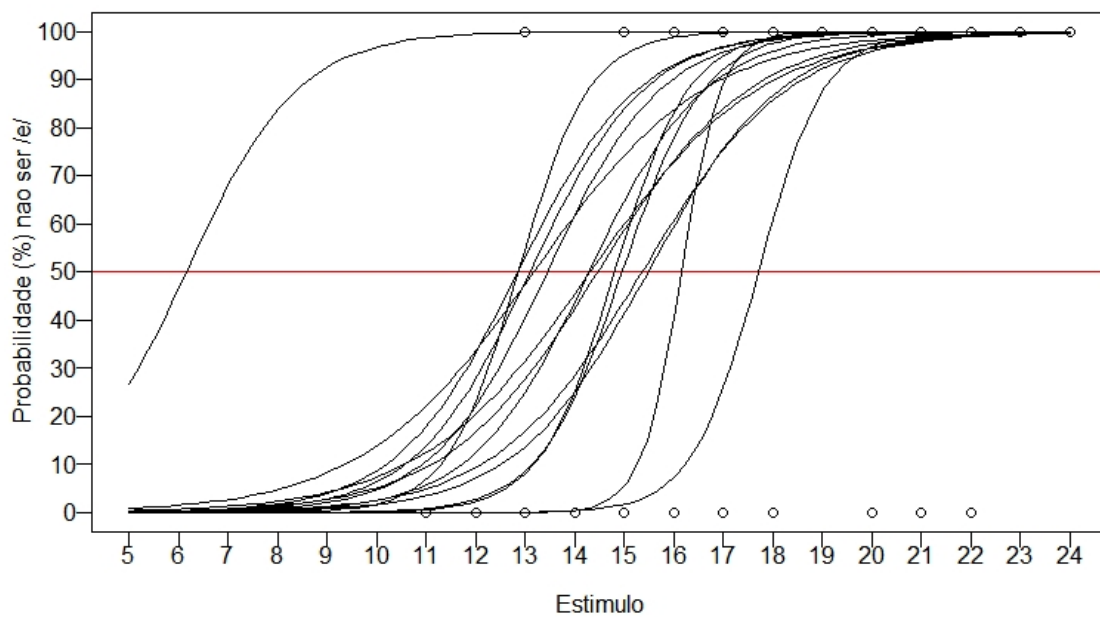


Figura 12: As curvas de regressão logística de todos os participantes para a transição da categoria /e/ para não /e/.

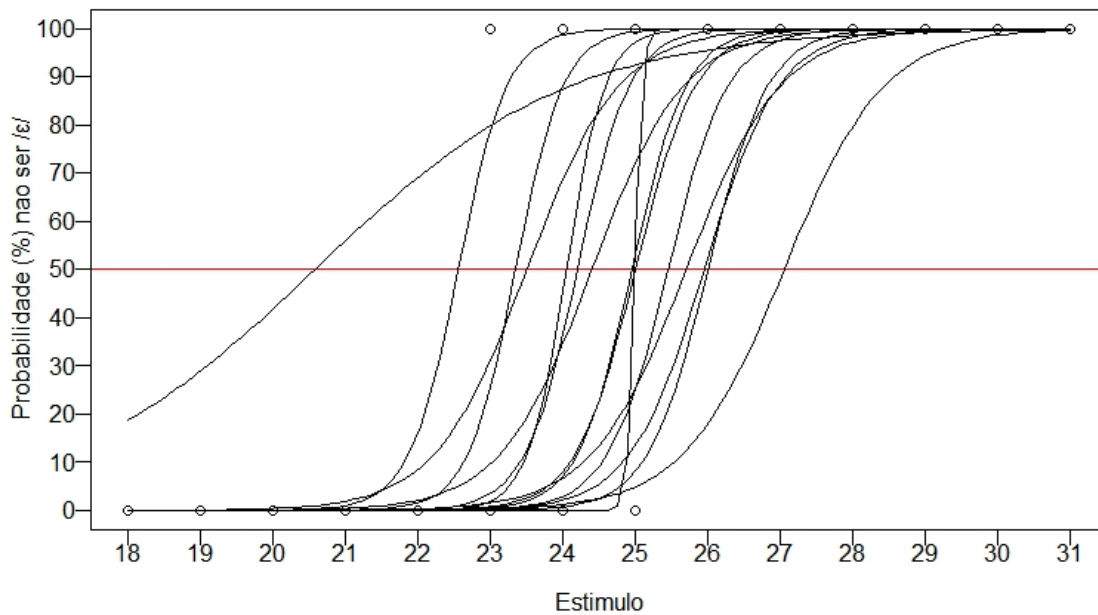


Figura 13: As curvas de regressão logística de todos os participantes para a transição da categoria /ε/ para não /ε/.

A partir das curvas de regressão logística ajustadas para todos os participantes (cf. Figuras 11, 12 e 13) obtiveram-se as estimativas (valores médios e medianas, com o desvio padrão entre parêntesis) das três fronteiras: /i/-/e/ (doravante fronteira 1), /e/-/ε/ (fronteira 2), /ε/-/a/ (fronteira 3). Na última coluna da Tabela 8 encontram-se os valores dos formantes dos estímulos que correspondem a cada fronteira.

Tabela 8: Fronteiras médias estimadas entre categorias e os valores dos formantes correspondentes

	Fronteiras médias (\pm desvio padrão)	Fronteiras medianas	Valores dos formantes ⁴⁹	
/i/-/e/ (fronteira 1)	5.28 (\pm 2.40)	5.65	F1:	344 Hz
			F2:	2049 Hz
			F3:	2754 Hz
/e/-/ε/ (fronteira 2)	14.61 (\pm 1.42)	14.46	F1:	406 Hz
			F2:	1893 Hz
			F3:	2503 Hz
/ε/-/a/ (fronteira 3)	24.52 (\pm 1.59)	24.95	F1:	551 Hz
			F2:	1648 Hz
			F3:	2338 Hz

⁴⁹ Os valores dos formantes correspondem aos valores dos estímulos mais próximos dos medianos. Assim sendo, a fronteira 1 \cong 6, a fronteira 2 \cong 15 e a fronteira 3 \cong 25.

A distribuição destes valores está representada no diagrama de extremos e quartis – cf. Figura 14. A partir desta representação, observa-se uma clara distinção entre as três diferentes fronteiras. É de salientar a reduzida variabilidade dos três grupos, em particular, do grupo correspondente à fronteira 2, que não exhibe *outliers*⁵⁰. Relativamente a fronteira 1, podemos confirmar a existência de dois participantes que se comportam de forma bastante diferente do resto do grupo. No que toca à fronteira 3, existe apenas um participante nestas condições.

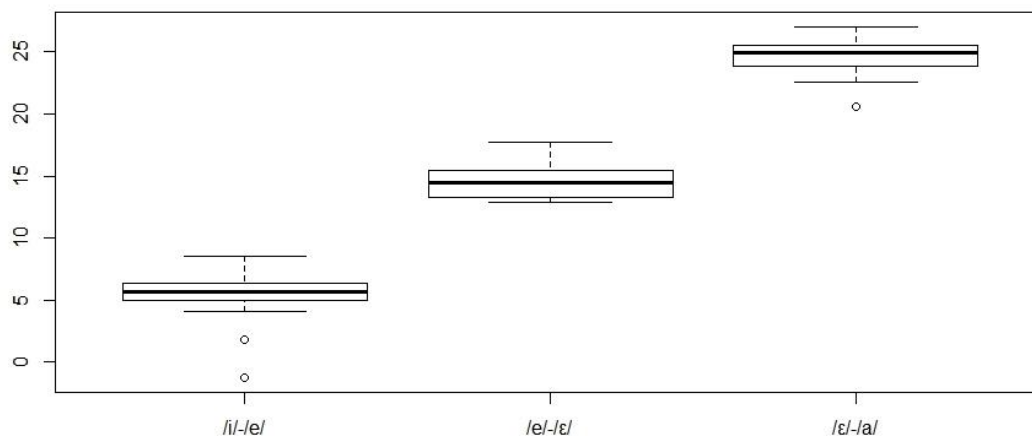


Figura 14: A distribuição dos valores das três fronteiras, /i/-/e/, /e/-/ɛ/ e /ɛ/-/a/ para os 15 participantes.

5.2. Fronteiras entre categorias

De modo a verificar se os participantes reconheceram a diferença entre categorias, foi feita uma análise estatística, de modo a confirmar se há ou não uma distância significativa entre as amostras das três fronteiras marcadas por cada indivíduo. Inicialmente, realizou-se o teste de normalidade (Shapiro-Wilk) para decidir entre uma abordagem paramétrica ou não paramétrica. Este teste foi escolhido devido ao reduzido número de observações por grupo (15) na nossa experiência. Para a amostra da fronteira 1 rejeita-se a hipótese de normalidade (valor $p < 0.05$), para a da fronteira 2 e 3 não se rejeita a hipótese de normalidade (valor $p > 0.05$, em ambos os casos). Recorreu-se, ainda, a dois adicionais testes de normalidade: Kolmogorov-Smirnov (KS)

⁵⁰ Relativamente às duas observações *outliers* do primeiro conjunto de fronteiras, mesmo que fossem excluídas teríamos: para a fronteira /i/-/e/, mediana = 5.70, /e/-/ɛ/ = 14.46 e /ɛ/-/a/ = 24.95, o que não difere significativamente daquilo que é obtido com a inclusão desses dois valores (-1.26 e 1.85) -ver Tabela 11. Optamos, portanto, por mantê-las na nossa análise. Além disso, estes dados forneceram-nos uma informação interessante relativa à categoria /i/ que discutiremos adiante.

com correção de Lilliefors e Anderson-Darling (AD), para verificar os resultados obtidos com o teste de Shapiro-Wilk. Tanto com o teste de normalidade de Kolmogorov-Smirnov (KS) com correção de Lilliefors como com o teste de Anderson-Darling (AD) obtiveram-se os seguintes resultados: para a fronteira 1 rejeita-se a hipótese de normalidade (valor $p < 0.05$), para as fronteiras 2 e 3 não se rejeita a hipótese de normalidade (valor $p > 0.05$, em ambos os casos). Os resultados dos 3 testes podem ser observados na Tabela 9, onde se apresentam os respectivos valores p .

Tabela 9: Valores de p dos testes de normalidade para as diferentes amostras

Amostra correspondente a:	Shapiro-Wilk	Kolmogorov-Smirnov (KS) com correção de Lilliefors	Anderson-Darling (AD)
Fronteira 1	0.0242*	0.0476*	0.0215*
Fronteira 2	0.4090	0.7251	0.5646
Fronteira 3	0.4995	0.5935	0.5097

*Significa a rejeição da hipótese de normalidade da distribuição.

Apesar de, para duas das três amostras, não se ter rejeitado a hipótese de distribuição normal, uma vez que se usa sempre os três grupos para fazer comparações, foi necessário optar por testes de hipóteses que não assumam a normalidade, pois esta teria que ser assumida para todos os grupos para ser possível compará-los num contexto paramétrico. Com o teste não paramétrico de Kruskal-Wallis (ou ANOVA não paramétrica), a hipótese nula da igualdade de medianas entre grupos foi rejeitada ($X^2=39.133$, $df = 2$, valor $p = 3.18 \times 10^{-9}$). Os testes post-hoc (o teste de Nemenyi com duas alternativas para o cálculo de valores p , sendo a primeira não corrigida para o caso de observações com o mesmo *rank* e a segunda corrigida para o caso de observações com o mesmo *rank*) resultaram em conclusões idênticas, ou seja, a comparação entre dois quaisquer grupos resulta num valor p menor que 0.05, quer se use a estatística de teste de Tukey ou do Qui-quadrado. As diferenças entre os grupos podem então ser consideradas estatisticamente significativas como se pode observar na Tabela 10.

Tabela 10: Resultado das comparações múltiplas entre grupos (perspetiva não paramétrica e paramétrica)

Comparações entre os grupos	Valor p do teste de Nemenyi com a estatística de teste de:		IC para a diferença entre as médias dos grupos, de acordo com o teste HSD de Tukey
	Tukey	Qui-quadrado	
Fronteira 1-2	0.005*	0.0075*	(-10.97; -7.68)
Fronteira 1-3	1.2×10^{-9} *	3.2×10^{-9} *	(-20.89; -17.60)
Fronteira 2-3	0.005*	0.0075*	(-11.56; -8.27)

*Significa a existência de diferenças significativas entre dois grupos, ao nível de significância de 5%.

Os procedimentos estatísticos paramétricos são mais robustos do que os não paramétricos. Assim, foi também realizada uma ANOVA paramétrica, supondo-se que a normalidade teria sido validada (estatística $F=403.7$, $df=2$, valor $p < 2 \times 10^{-16}$). Tendo-se obtido os mesmos resultados no contexto paramétrico, são sustentáveis os resultados anteriores. Também no contexto paramétrico foram realizadas comparações múltiplas de acordo com o teste HSD de Tukey, cujo resultado se apresenta na última coluna da Tabela 10. Nenhum dos intervalos de confiança a 95 % para a média das diferenças entre grupos contém o zero, pelo que é de esperar que as diferenças entre médias possam ser significativamente diferentes.

Tendo em conta estes resultados estatísticos confirmaram-se quatro categorias vocálicas divididas com três fronteiras significativamente diferentes.

5.3. Tarefa de classificação de *goodness*

Tal como explicado na secção 4.6. *Procedimento e design da experiência*, após a tarefa de identificação os participantes classificaram a qualidade do som ouvido numa escala de 1 a 5 conforme a vogal escolhida na tarefa de identificação. Assim sendo, um estímulo pode receber no máximo 100 pontos por cada participante, caso este seja classificado 20 vezes como o melhor exemplo (5 pontos). Os pontos somados da classificação de *goodness* (de 1 a 5) de cada estímulo foram calculados para cada participante e representados num gráfico de bolhas – cf.

Anexo 5. Para a representação do grupo (todos os participantes), utilizou-se o valor mediano em vez do médio – cf. Figura 15, na media em que o resultado do teste da normalidade mostrou que o grupo de observações da fronteira entre /i/ e /e/ não segue uma distribuição normal– cf. Tabela 9, e nesse caso, a mediana deverá ser a medida tendência central que melhor representa os dados.

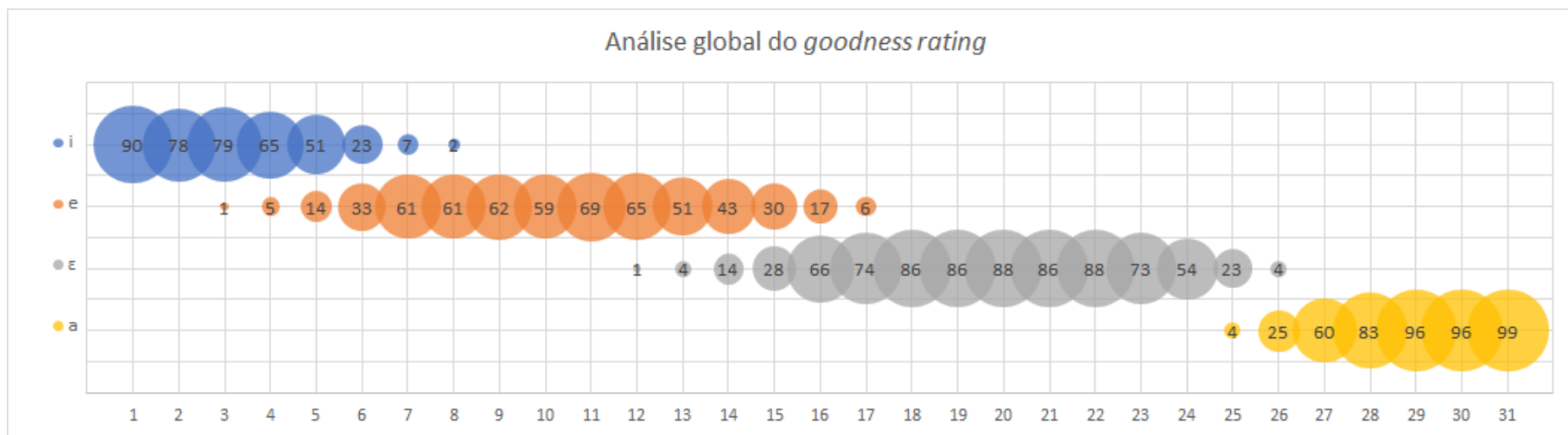


Figura 15: Representação do goodness rating mediano de todos os participantes, para cada um dos 31 estímulos (eixo horizontal). As diferenças de pontuação são representadas pelos tamanhos dos círculos. O número dentro de cada círculo é a pontuação mediana para cada estímulo, considerando todos os participantes, podendo variar de 1 (pior) a 100 (melhor). As cores dos círculos correspondem da seguinte forma às categorias que representam: azul - /i/, laranja - /e/, cinza - /ɛ/ e amarelo - /a/.

5.3.1. Estrutura interna de categorias e protótipos

A segunda questão do nosso estudo foi a verificação da existência da estrutura interna gradual da categoria, verificando, também, um som (ou sons) representativo(s) de cada categoria fonêmica, o protótipo, para os falantes nativos de PE. Conforme se constata na Figura 15, confirma-se uma estrutura de categoria em que o protótipo fica no centroide, com a pontuação mais alta. À medida que o estímulo se afasta do centroide a pontuação de *goodness* vai diminuindo. Salienta-se que nem todos os participantes responderam da mesma maneira, isto é, observou-se uma grande variabilidade individual na identificação da categoria dos estímulos e, conseqüentemente, uma grande variabilidade nos protótipos individuais, como se constata, por exemplo, nas Figuras 16 e 17. No entanto, os informantes, em geral, usaram quatro categorias com três fronteiras, exibindo também uma estrutura interna de cada categoria, com o seu protótipo.

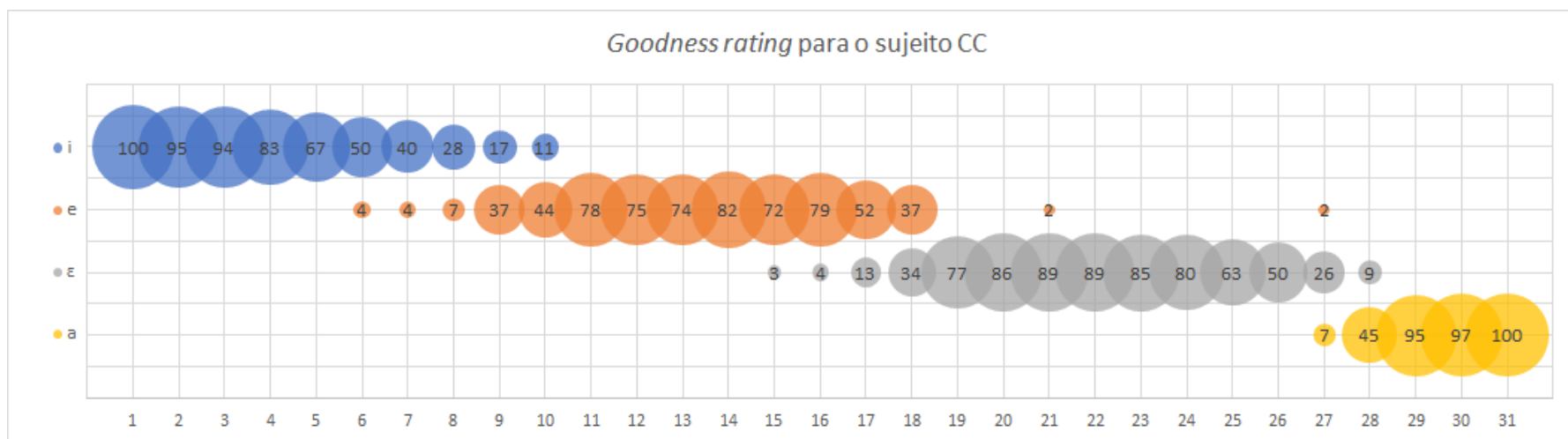


Figura 16: Representação do goodness rating do informante CC

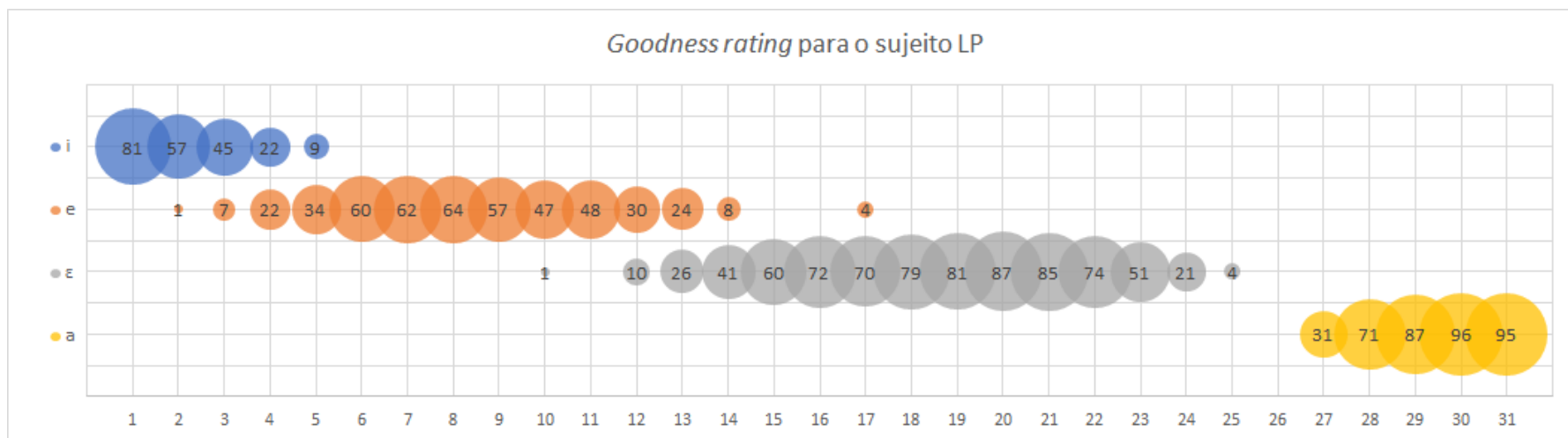


Figura 17: Representação do goodness rating do informante LP

Houve dois informantes (BM e AC) que tiveram um desempenho muito diferente dos outros na identificação da categoria /i/, referidos também no resultado da tarefa de identificação – cf. Rodapé 50. Estes dados, contudo, não foram excluídos, dado que se verificou o padrão de estrutura interna noutras categorias – cf. Figuras 18 e 19. Estes dados levam-nos a conjecturar, para estes informantes, uma localização mais periférica do protótipo de /i/. Discutiremos esta questão na secção 6.4. *Perceptual hyperspace* (Johnson, 2000) e *adaptive dispersion* (Liljencrants & Lindblom, 1972).

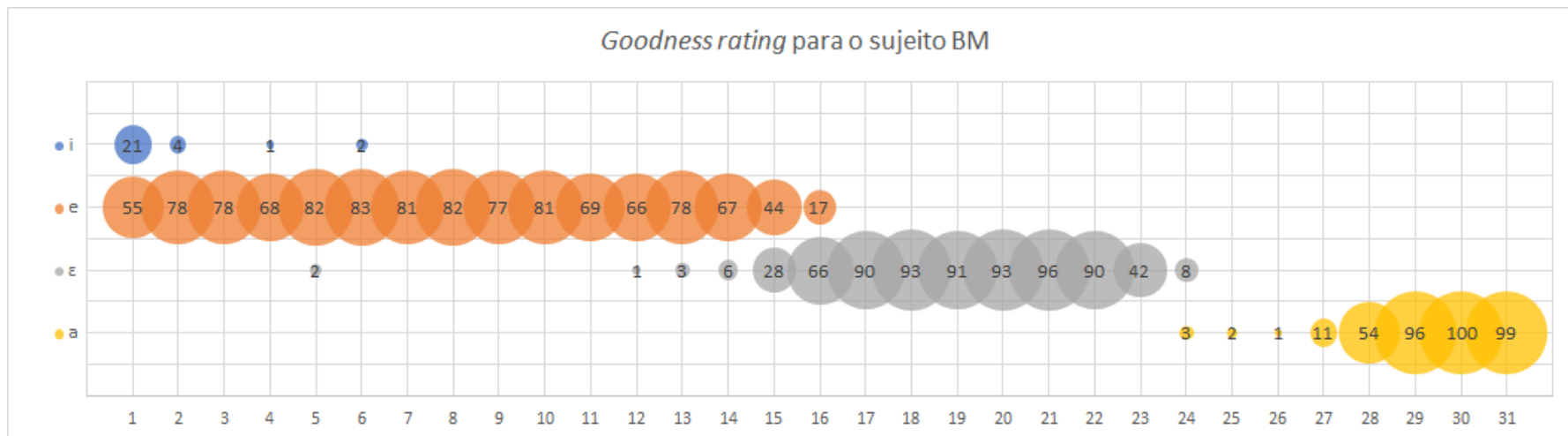


Figura 18: Representação do goodness rating do informante BM

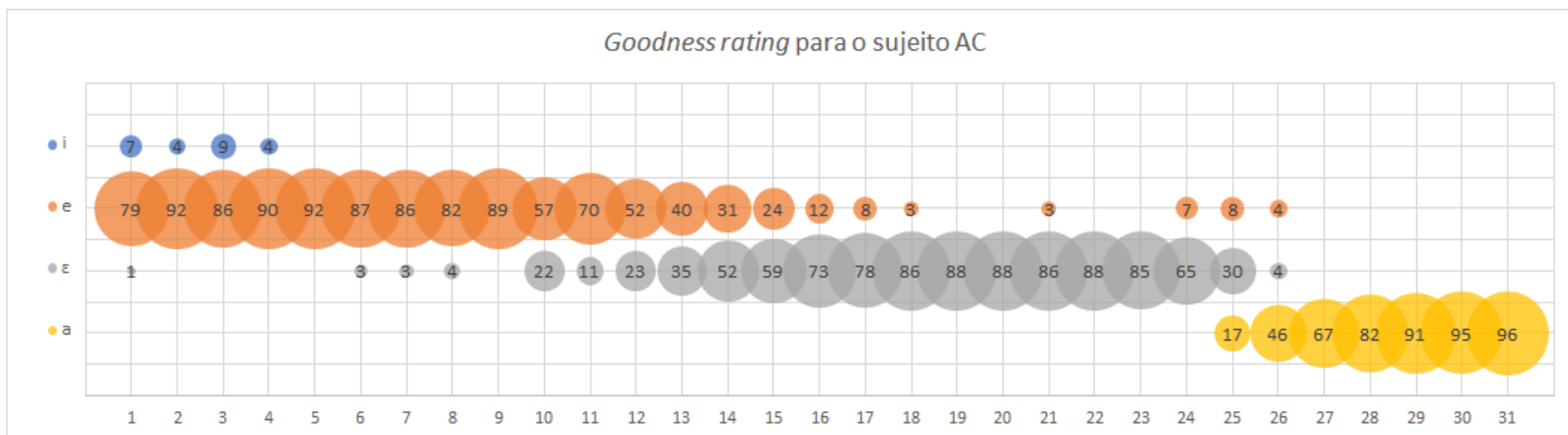


Figura 19: Representação do goodness rating do informante AC

Destaca-se ainda que o informante CL parece mostrar alguma insegurança na classificação de *goodness* para as categorias de /e/ e /ɛ/, mas não para as de /i/ e /a/– cf. Figura 20. Embora seja o único informante que se tenha comportado deste modo, não deixa de ser interessante conjecturar sobre uma característica destas categorias na percepção, que causasse este fenómeno.

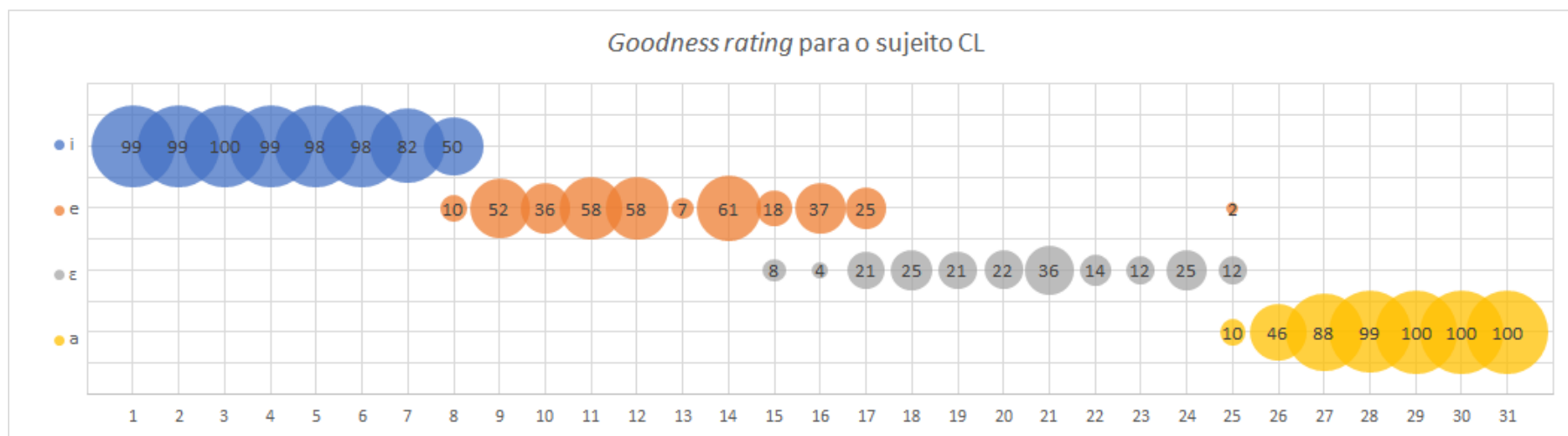


Figura 20: Representação do goodness rating do informante CL

5.3.2. Comparação de grau de nitidez entre as fronteiras

Tendo em conta as três fronteiras verificadas nos resultados de 5.1. *Tarefa de identificação* e 5.2. *Fronteiras entre categorias*, e a pontuação de *goodness* de 5.3, trata-se finalmente a questão principal do nosso estudo, isto é, se a fronteira entre /e/ e /ɛ/ é menos nítida do que outras fronteiras na percepção dos falantes nativos de PE. Antes de analisar estatisticamente as fronteiras com a pontuação de *goodness*, analisaram-se as respostas X. Salienta-se que, conforme se constata na Figura 21, o valor que representa a frequência absoluta (n) de respostas X para o estímulo que

corresponde à fronteira 3 é bastante maior (estímulo 25, $n = 86$), do que os que correspondem à fronteira 1 (estímulo 6, $n = 10$) ou à fronteira 2 (estímulo 15, $n = 20$). Optando neste estudo por interpretar a resposta X como uma decisão firme dos participantes, isto é, o som ouvido não é ambíguo, só não pertence a nenhuma das categorias apresentadas, podemos supor que, enquanto nas fronteiras 1 e 2, os participantes identificaram os sons ouvidos, selecionando uma das categorias, na fronteira 3, de facto, responderam mais frequentemente que os sons não pertencem a nenhuma. Ou seja, os participantes terão mais facilidade na distinção entre as categorias /ε/ e /a/, tendo selecionado mais vezes a resposta X.

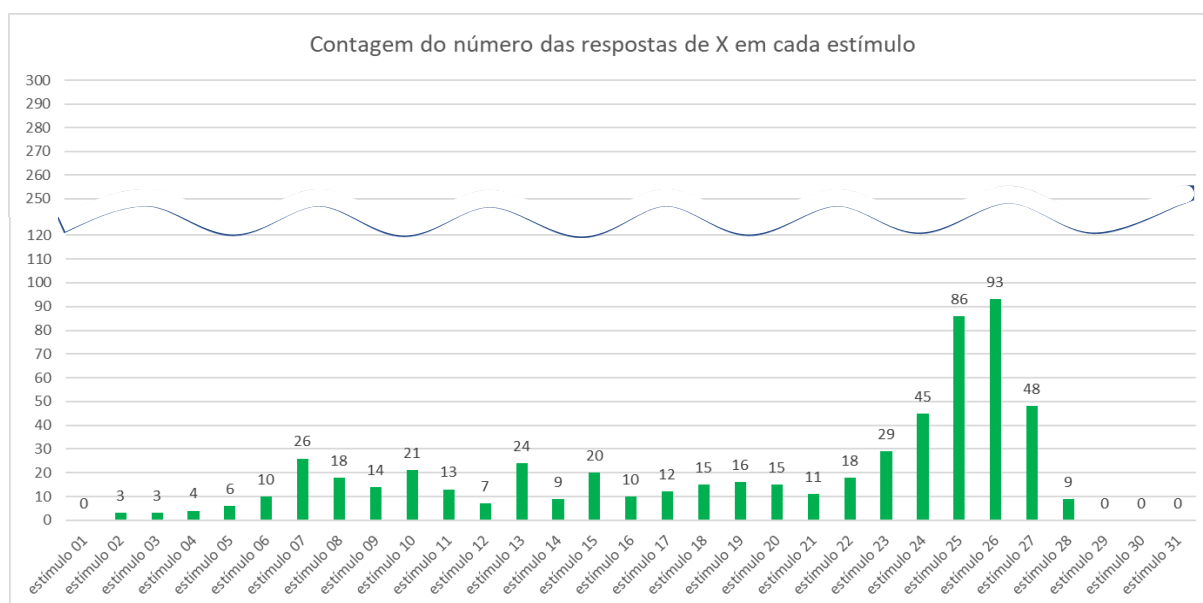


Figura 21: Os valores representados correspondem à frequência absoluta de respostas X para cada um dos 31 estímulos que se encontram no eixo horizontal. (A “onda” representa a omissão dos valores entre 120 e 250 no eixo vertical)

Supõe-se a existência de um conflito, entre duas categorias, aquando do processo de categorização. Poder-se-á, no sentido de o resolver, comparar, relativamente ao mesmo estímulo, os valores elevados de *goodness score* das duas categorias problemáticas. Partindo desta suposição, sujeitaram-se, a uma análise estatística, os pontos das duas categorias sobrepostas em cada fronteira. Ou seja, se entre os pontos de duas categorias sobrepostas não existir uma diferença significativa, poderia, então, haver uma fronteira menos nítida. Foram escolhidos os valores das fronteiras individuais, e não medianos, para esta análise, dado que se observou uma grande variabilidade entre indivíduos, em termos destes valores (os valores das fronteiras individuais podem até ser muito afastados dos valores médios ou medianos das

fronteiras⁵¹). Obtiveram-se, através das respostas da tarefa de identificação, as três fronteiras individuais entre categorias – cf. Tabela 11.

Tabela 11: Os valores individuais e os medianos das fronteiras. Os medianos são: /i/-e/=5.65, /e/-/ε/=14.46, /ε/-a/=24.95.

Participante	/i/-e/	/e/-/ε/	/ε/-a/
CC	9 (8.5 ≅ 9)	18 (17.70)	27 (27.05)
AP	6 (5.65)	13 (13.09)	24 (24.05)
BM	-1 (-1.26)	15 (14.95)	23 (23.34)
MM	2 (1.85)	13 (13.44)	24 (23.59)
MF	6 (5.65)	14 (14.28)	24 (24.40)
ML	7 (6.65)	15 (14.80)	23 (22.55)
CP	5 (5.20)	13 (12.84)	25 (25.00)
IM	5 (5.40)	15 (15.35)	26 (25.45)
MC	7 (6.75)	16 (16.15)	26 (26.00)
GS	6 (6.14)	16 (16.24)	25 (24.99)
FM	6 (6.05)	16 (15.47)	26 (25.70)
AC	5 (4.68)	13 (13.16)	25 (24.95)
BR	6 (5.70)	15 (14.46)	26 (25.95)
CL	8 (8.13)	14 (14.31)	21 (20.57)
LP	4 (4.05)	13 (12.85)	24 (24.20)
Mediana	6 (5.65)	15 (14.46)	25 (24.95)

Analísaram-se, primeiro, a normalidade da distribuição dos pontos de *goodness* de cada categoria, de modo a escolher um teste estatístico adequado. Sendo o número de observações reduzido (n=15), foi escolhido o teste de normalidade de Shapiro-Wilk. Relativamente à fronteira 1, para a amostra dos pontos de *goodness* de /i/ não se rejeita a hipótese de normalidade (valor $p = 0.6725 > 0.05$) e para a amostra dos pontos de *goodness* de /e/ rejeita-se a hipótese de normalidade (valor $p = 0.0255 < 0.05$). Na fronteira 2, para a amostra dos pontos de *goodness* de /e/ não se rejeita a hipótese de normalidade (valor $p = 0.9732 > 0.05$) e para a amostra dos pontos de *goodness* de /ε/ não se rejeita a hipótese de normalidade (valor

⁵¹ A variabilidade individual relativa às fronteiras medianas pode também ser devido à dimensão reduzida da amostra (apenas 15 participantes).

$p = 0.3635 > 0,05$). Na fronteira 3, para a amostra dos pontos de *goodness* de /ε/ não se rejeita a hipótese de normalidade (valor $p = 0.6445 > 0.05$), e para a amostra dos pontos de *goodness* de /a/ na fronteira /ε/-/a/ rejeita-se a hipótese de normalidade (valor $p = 0.0048 < 0.05$). Recorreu-se, ainda, a dois outros testes de normalidade, o teste de Kolmogorov-Smirnov (KS) com correção de Lilliefors e o teste de Anderson-Darling (AD) – cf. Tabela 12. É de notar a discordância entre os testes de KS com correção de Lilliefors e Shapiro-Wilk, sendo o primeiro mais permissivo do que o segundo. O teste de AD, por sua vez, está em concordância com o teste de KS com correção de Lilliefors, mas pode-se dizer que está na fronteira entre o de KS e o de Shapiro-Wilk, o menos permissivo dos três. No entanto, como as amostras têm todas 15 observações, dimensão reduzida, opta-se por agir em conformidade com o resultado do teste de Shapiro-Wilk.

Uma vez que, para a fronteira 2, os pontos de *goodness* das duas categorias não põem de parte a hipótese de haver uma distribuição normal subjacente aos dados, a sua análise pode ser feita com o teste *t* para amostras emparelhadas. Relativamente aos pontos de *goodness* da categoria /e/ na fronteira 1 e da categoria /a/ na fronteira 3, no entanto, rejeitou-se a possibilidade de haver uma distribuição normal. Logo, para estas duas fronteiras foi escolhido o teste de Wilcoxon para a sua análise. É de notar que tanto o teste *t* para amostras emparelhadas como o teste de Wilcoxon são adequados à análise pretendida, uma vez que se tem como objetivo comparar observações da mesma variável dependente (valor do *goodness score*) em duas condições diferentes (duas categorias) mas numa perspectiva intra-sujeito, isto é, de amostras emparelhadas.

Tabela 12: Valores de *p* dos testes de normalidade para as diferentes amostras

Amostras de goodness em:		Shapiro-Wilk	Kolmogorov-Smirnov (KS) com correção de Lilliefors	Anderson-Darling (AD)
Fronteira 1	/i/	0.6725	0.5040	0.3805
	/e/	0.0255*	0.5028	0.0522
Fronteira 2	/e/	0.9732	0.8688	0.6851
	/ε/	0.3635	0.3975	0.2617
Fronteira 3	/ε/	0.6445	0.5214	0.8145
	/a/	0.0048*	0.0129*	0.0013 *

*Significa a rejeição da hipótese de normalidade da distribuição.

Assim sendo, relativamente à fronteira 1, realizou-se o teste de Wilcoxon para comparação de medianas entre amostras emparelhadas, colocando a hipótese nula (H_0): não há diferença entre as medianas dos pontos de *goodness* das duas categorias que se encontram numa dada fronteira; e a hipótese alternativa (H_1): existe uma diferença entre as medianas dos pontos de *goodness* das duas categorias que se encontram numa dada fronteira. Dos resultados apresentados na Tabela 13, pode-se concluir que não há diferenças significativas entre os valores medianos (valor $p=0.5760 > 0.05$) na fronteira 1. Relativamente à fronteira 2, optou-se por considerar o resultado do teste *t* para amostras emparelhadas que, sendo um teste paramétrico, é mais potente que o teste de Wilcoxon e havia evidências, atendendo aos três testes de normalidade, que a comparação de *goodness scores* na fronteira 2 podia ser realizada num contexto paramétrico. De acordo com o teste *t*, não há razão para considerar que as médias dos *goodness scores* sejam significativamente diferentes (valor $p=0.0929 > 0.05$). Relativamente à comparação dos *goodness scores* na fronteira 3, novamente considera-se o resultado do teste de Wilcoxon, e pode-se concluir que os *goodness scores* medianos podem ser considerados significativamente diferentes (valor $p = 0.0170 < 0.05$) – cf. Tabela 13.

Tabela 13: Testes de comparação de medianas/ médias para amostras emparelhadas

	Resultado dos testes:			
	Valor <i>p</i> do teste de Wilcoxon	Valor <i>p</i> do teste <i>t</i>	IC 95% (dif. médias)	Média das diferenças
Fronteira 1	0.5760	0.9209	(-18.8786, 20.7358)	0.9286
Fronteira 2	0.0499*	0.0929	(-1.7035, 19.7035)	9.0000
Fronteira 3	0.0170*	0.0107*	(4.7537, 30.3130)	17.5333

*Significa existências de diferenças significativas entre grupos ao nível de significância de 5 %

A comparação do grau de nitidez das fronteiras foi também analisada empiricamente, recorrendo ao cálculo do rácio entre os *goodness scores* medianos nas fronteiras entre categorias, considerando fronteira 1= 6, fronteira 2=15 e fronteira 3=25. Volta-se a referir a nossa suposição: no processo da categorização do *continuum*, deve haver um conflito na fronteira entre duas categorias. Quanto mais próximo de 1 for o rácio entre os *goodness score* numa fronteira, mais confusa será a decisão de categorização. Conforme se constata na tabela 14, comparando as três fronteiras, conclui-se que a fronteira 2 é a que apresenta o rácio mais próximo de 1, o que sugere um maior grau de confusão. Em segundo lugar, surge a fronteira 1,

com o rácio de 0.70, e por fim a fronteira 3 com o rácio de 0.17, o que sugere menor grau de confusão que as duas anteriores.

Tabela 14: Resultado do cálculo do rácio entre goodness scores nas três fronteiras

	Nº de estímulo	Rácio
Fronteira 1	6	23/33= 0.70
Fronteira 2	15	28/30= 0.93
Fronteira 3	25	4/23= 0.17

Tendo em conta as várias análises realizadas, o grau de nitidez da fronteira 3 pode ser considerado o maior. É o que corresponde ao menor valor do rácio, e é também a única fronteira relativamente à qual se pode considerar que existem diferenças significativas entre as medianas dos *goodness scores*. De acordo com a análise estatística, existe um grau de confusão considerável nas duas primeiras fronteiras, uma vez que não há diferenças significativas entre os seus valores médios/medianos dos *goodness scores*. Podemos, no entanto, referir que a média das diferenças entre grupos é maior no caso da fronteira 2 do que no caso da fronteira 1 (ver a Tabela 13), o que pode sugerir uma maior nitidez na fronteira 2. Por outro lado, de acordo com a análise empírica do cálculo do rácio, a fronteira 2 será a menos nítida das três.

Capítulo 6. – Discussão dos resultados

Neste capítulo discutimos os resultados obtidos no nosso estudo, tendo em conta as hipóteses formuladas inicialmente. Verificamos, ainda, se o objetivo principal do nosso estudo foi cumprido, isto é, averiguamos se existe menor clareza da distinção entre as vogais médias anteriores na percepção dos nativos do PE, devido à complexidade das relações que estas vogais estabelecem no seu sistema.

As hipóteses formuladas na *Introdução* são as seguintes:

1. Existência de quatro categorias fonémicas com uma estrutura gradual (embora possa haver dispersão intra e inter-ouvintes tal como confirmado na produção);
2. O grau da nitidez da fronteira entre as vogais médias /e/ e /ɛ/ é menor do que outras fronteiras.

Além disso, procuramos relacionar os nossos resultados com os resultados de estudos e teorias anteriores, referidos nos capítulos 2, *Influência do sistema fonológico e as vogais do PE* e 3, *Os estudos experimentais sobre a influência da relação fonológica na percepção da fala*, de forma a fornecer evidências experimentais.

6.1. As quatro categorias vocálicas

A primeira hipótese do presente estudo foi concretizada separadamente usando as duas questões seguintes: 1) se se encontram quatro categorias fonémicas; 2) se estas categorias possuem uma estrutura interna gradual. Relativamente à primeira questão, através dos dados obtidos na tarefa de identificação, verificaram-se quatro categorias vocálicas marcadas com três fronteiras estatisticamente distintas. Estes resultados mostraram que os falantes nativos de PE (ou pelo menos os falantes jovens de Lisboa) foram capazes de distinguir quatro categorias vocálicas nos sons utilizados nesta experiência. Assim sendo, presume-se que os falantes nativos de PE possuem estas quatro categorias na sua representação mental. Estes resultados correspondem às descrições fonológicas do sistema vocálico do PE, em que se encontram quatro fonemas não recuados: /i, e, ɛ, a/. A conformidade com as descrições fonológicas foi

também relatada nos dados da produção obtidos por Escudero *et al.* (2009: 1385)⁵², cujos participantes possuíam as mesmas características dos do nosso estudo. Assim sendo, relativamente a esta questão, a descrição fonológica, a produção e a percepção estão em concordância. Além disso, tal como relatado no estudo de Aaltonen *et al.* (1997), confirmou-se uma grande variabilidade individual no local onde as fronteiras se encontram. De acordo com Repp e Liberman (1987: 32), além dos vários fatores metodológicos⁵³, a localização da fronteira entre categorias fonéticas no estudo da percepção da fala pode ser influenciada pelos fatores individuais como uma representação mental distinta das categorias e/ou uma estratégia diferente aquando da decisão de categorização do indivíduo (cf. Aaltonen *et al.*, 1997: 1093).

6.2. Estrutura interna da categoria vocálica

Relativamente à segunda questão, confirmou-se, através da tarefa da classificação de *goodness*, uma estrutura interna das categorias vocálicas, em que o estímulo com a pontuação mais alta, considerado como o protótipo dessa categoria, fica aproximadamente no centroide e a pontuação de *goodness* vai diminuindo à medida que o estímulo se afasta do centroide. Os resultados obtidos através do nosso estudo vão de encontro aos resultados dos estudos anteriores em que se confirmou uma estrutura interna da categoria fonética⁵⁴ (Grieser & Kuhl, 1989; Kuhl, 1991; Miller, 1994). Em geral, os protótipos foram encontrados, em termos do número de

⁵² Em Escudero *et al.* (2009: 1385), recorrendo ao cálculo estatístico de análise de variância com medidas repetidas, verificou-se o principal efeito de F₁ na categoria vocálica de ambas as variantes do português. Este facto está em concordância com a proposta de Veloso (2016) para PE e de Wetzels (1992) para PB, referida na secção 2.3. *Propostas teóricas do inventário fonológico das vogais do PE*, isto é, as vogais fonológicas do português opõem-se pelos 4 graus de abertura. No presente estudo, não se realizou uma análise detalhada para chegar à mesma conclusão, limitando, apenas, presumir a existência das 4 categorias distintas na representação mental dos falantes do dialeto em questão.

⁵³ Os autores deram exemplos de fatores metodológicos que poderão influenciar o local da fronteira, como o contexto fonético, velocidade da fala, mistura de pistas acústicas e a experiência linguística (Repp & Liberman, 1987: 33)

⁵⁴ A verificação da sua teoria de “*perceptual magnet effect*” (Iverson & Kuhl, 1995; Kuhl *et al.*, 1992; Kuhl, 1991) não se enquadrava nos objetivos deste estudo.

estímulos, mais ou menos no centro da categoria identificada pelos participantes⁵⁵. No entanto, enquanto na experiência de Kuhl (1991) se obtiveram respostas consistentes em todos os participantes, os nossos dados mostram variabilidade individual dos protótipos de cada categoria. A variação individual do protótipo tinha sido observada antes por Lively e Pisoni (1997) e Aaltonen *et al.* (1997).⁵⁶ Seria interessante realizar-se uma comparação entre a variabilidade individual da produção e do protótipo, para averiguar a existência de relações entre estes.

6.3. Grau de nitidez das fronteiras

De modo a testar a nossa segunda hipótese, isto é, verificar se a fronteira entre as vogais médias apresenta uma menor nitidez na percepção dos falantes jovens nativos de Lisboa, realizou-se uma análise estatística com os dados da tarefa de classificação de *goodness*. Esta hipótese foi sustentada pela complexidade das relações que estas vogais estabelecem entre as formas subjacentes e as suas realizações fonéticas (Veloso, 2016), apontada como a provável dificuldade na aquisição destas relações fonológicas pelas crianças (Freitas, 2004) e a observação da convergência entre as vogais médias na produção dos falantes deste dialeto (Escudero *et al.*, 2009). Foram analisadas as diferenças entre os pontos de *goodness* das duas categorias de cada fronteira, de modo a avaliar o grau de confusão na decisão de categorização, empregando as fronteiras individuais marcadas na tarefa de identificação. Deste modo, revelou-se, consistentemente, uma maior nitidez da fronteira 3 (/ε-/a/) do que das outras duas fronteiras, uma vez que as análises da frequência absoluta (n) de respostas X, a estatística e o cálculo do

⁵⁵ Na literatura relativa à percepção da fala existem duas propostas diferentes referentes ao local onde se localiza o protótipo: o local mais periférico da categoria, por exemplo, a teoria da dispersão (Johnson *et al.*, 1993; Johnson, 2000) ou no centroide (Kuhl, 1992). Os resultados obtidos em Eerola *et al.* (2014) sustentam a teoria de Johnson *et al.* (1993) e Johnson, (2000), verificando-se que o local do protótipo perceptivo é mais periférico do que a média aritmética da categoria. Como os estímulos utilizados no nosso estudo não cobrem todo o espaço de uma categoria, não podemos responder à questão sobre o local do protótipo. Assim sendo, fazemos apenas uma observação relativa à existência da estrutura interna que inclui o protótipo, cuja pontuação é a mais alta, e que se encontra, em termos de número de estímulos, aproximadamente no centro das categorias /e/ e /ε/.

⁵⁶ No estudo referido acima (Eerola *et al.*, 2014), empregando o *goodness rating*, as três medidas diferentes dos protótipos foram avaliadas: o protótipo absoluto; o centroide do protótipo e o protótipo ponderado, de modo a estruturar a categoria interna vocálica do alemão e do finlandês. Confirmou-se, com a metodologia de protótipo absoluto, a especificidade de cada língua e uma maior variabilidade individual em termos do local do protótipo. Por outro lado, obteve-se, em ambas as línguas, menor variabilidade entre sujeitos e até entre línguas com a metodologia do protótipo ponderado.

rácio estão em concordância. Relativamente à análise da frequência absoluta de respostas X, o número relevante da resposta X na fronteira 3 (marcada pelo valor mediano de todos os participantes) podia ter surgido da inadequação da trajetória de /ɛ/ a /a/ consoante a produção real dos falantes nativos portugueses. Contudo, a mesma análise com os dados do pré-teste não demonstrou nenhum destaque nesta fronteira – cf. Figura 22. Assim sendo, o resultado das análises da frequência absoluta de respostas X mantém-se como o argumento que sustenta a possibilidade da maior nitidez da fronteira 3 na percepção dos participantes lisboetas.

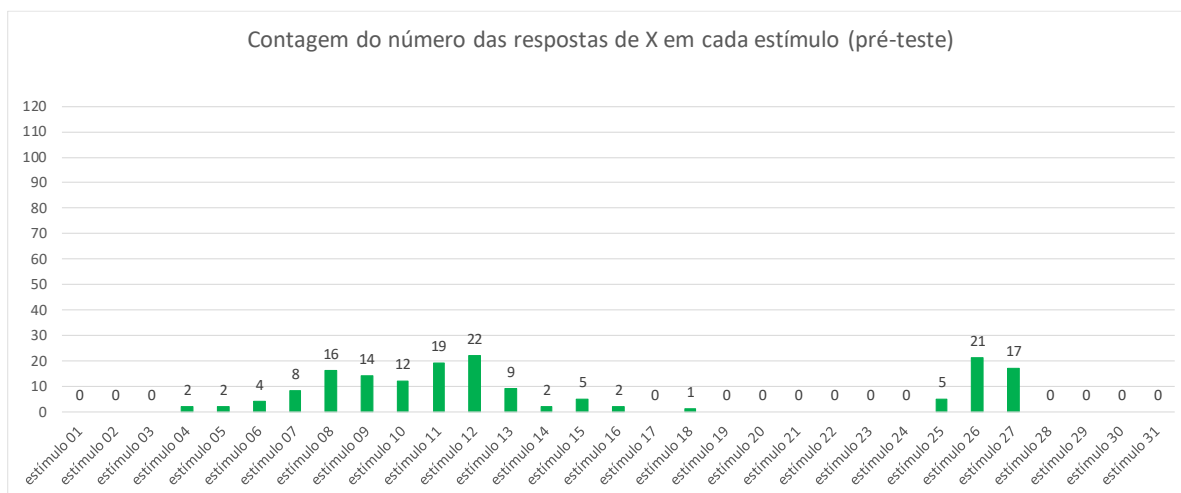


Figura 22: Os valores representados correspondem à frequência absoluta de respostas X no pré-teste para cada um dos 31 estímulos que se encontram no eixo horizontal.

No que concerne às outras duas fronteiras, os resultados não nos permitiram chegar a uma conclusão pois estes foram inconclusivos. Isto é, decidir qual das duas fronteiras apresenta uma maior nitidez está dependente do método utilizado para a análise. O estudo de Silva e Neves (2016), em que se estudou a categorização perceptiva das vogais médias anteriores do PB, verificou a existência de uma melhor distinção entre /i/-/e/ do que entre /e/ e /ɛ/. Os resultados estatísticos da nossa experiência mostraram uma maior nitidez da fronteira 2 em comparação com a fronteira 1, o que difere dos resultados de Silva e Neves (2016). Esta diferença pode dever-se aos seguintes fatores: 1) as diferentes especificidades das duas variantes; 2) as metodologias distintas utilizadas em cada experiência, e ainda, 3) as limitações da nossa experiência (este último veremos adiante). No que diz respeito ao primeiro fator, podemos considerar os estudos teóricos fonológicos e os da aquisição, sendo os últimos uma possível

prova para as propostas teóricas. Fikkert (2005) e Lee (2010) apresentaram uma proposta relativa à ordem de construção do sistema vocálico de cada variante, baseando-se na Teoria de Hierarquia Contrastiva (Dresher, 2003, 2004). De acordo com as suas propostas, a construção gradual das oposições vocálicas de duas variantes é a mesma⁵⁷. Fikkert (2005: 276) defende-a com os dados da produção de duas crianças nativas do PE, em que se observaram variações de alturas entre [i] e [ε]. Quanto às crianças nativas do PB, Bonilha (2004) confirmou uma aquisição mais tardia das vogais semiabertas⁵⁸ (Bonilha, 2004). O que difere entre as duas variantes é a representação fonética no nível de superfície. Ou seja, cada variante possui processos fonológicos distintos. Tal como referido no rodapé 17 da secção 2.3. *Propostas teóricas do inventário fonológico das vogais do PE*, o processo de neutralização das vogais médias é um dos exemplos desses processos (Bisol & Veloso, 2016: 70). Assim sendo, apesar das duas variantes do português apresentarem, aparentemente, o mesmo processo fonológico, os tipos distintos de neutralização das vogais médias podem afetar de modo diferente a percepção dos nativos de cada variante. Por outro lado, o fator da especificidade pode também ser dialetal. Embora o pré-teste com participantes do Norte de Portugal não tenha sido desenhado para fazer uma comparação com os resultados do teste principal, será interessante referir algumas observações. Em comparação com os dados dos lisboetas, constatam-se padrões distintos das sobreposições dos *goodness scores* entre categorias, na Figura 23 (que representa os *goodness scores* medianos dos nortenhos). A maior sobreposição é observada na transição entre as categorias /e/ e /ε/, estando 17 estímulos sobrepostos, enquanto, nos resultados dos falantes lisboetas, o número dos estímulos sobrepostos na transição entre as categorias /e/ e /ε/ é 6, sendo igual ao número dos estímulos sobrepostos entre /i/ e /e/. É pertinente referir os comentários dos participantes, recolhidos após cada experiência. Alguns participantes lisboetas mencionaram a dificuldade em distinguir /i/ de /e/. Por exemplo, o sujeito LP mencionou que a duração da vogal /i/ (na posição acentuada em português) nunca é tão longa como nos nossos estímulos, pelo que às vezes provocou uma sensação de estranheza⁵⁹. Os participantes do Norte

⁵⁷ Fikkert (2005) propôs os traços do ponto de articulação para as primeiras dicotomias, enquanto Lee (2010) utiliza os traços da posição da língua. Embora as propostas para definir os traços nas dicotomias sejam diferentes entre as duas variantes, a construção gradual é a mesma.

⁵⁸ Em comparação com a aquisição das consoantes, a das vogais é relativamente precoce. No PB, contudo, as crianças começam a adquirir, primeiro, as vogais periféricas: /a/ (1:0), e depois /i, u/ (1:1). Na segunda fase (1:2), aparecem as médias altas /e, o/, e finalmente, as médias baixas /ε/ (1:6), /ɔ/ (1:5), sendo /ε/ a mais tardia a ser adquirida (1:8). (Bonilha, 2004: 65-66)

⁵⁹ Apesar deste comentário, o desempenho do sujeito LP não mostrou nenhuma irregularidade- cf. Figura 17.

(pré-teste) não expressaram o mesmo tipo de opinião, referindo, por sua vez, a dificuldade em distinguir entre /e/ e /ɛ/. Estes comentários, que se verificaram, tendencialmente, diferentes, quer os sujeitos sejam oriundos do Norte ou de Lisboa, podem ser expressão de especificidades dialetais. Estamos conscientes de que o número de amostras dos participantes do Norte é muito reduzido (6) e, por isso, é necessário verificar esta hipótese, realizando um estudo comparativo da perceção da fala entre os dois dialetos com um número maior de amostras⁶⁰.

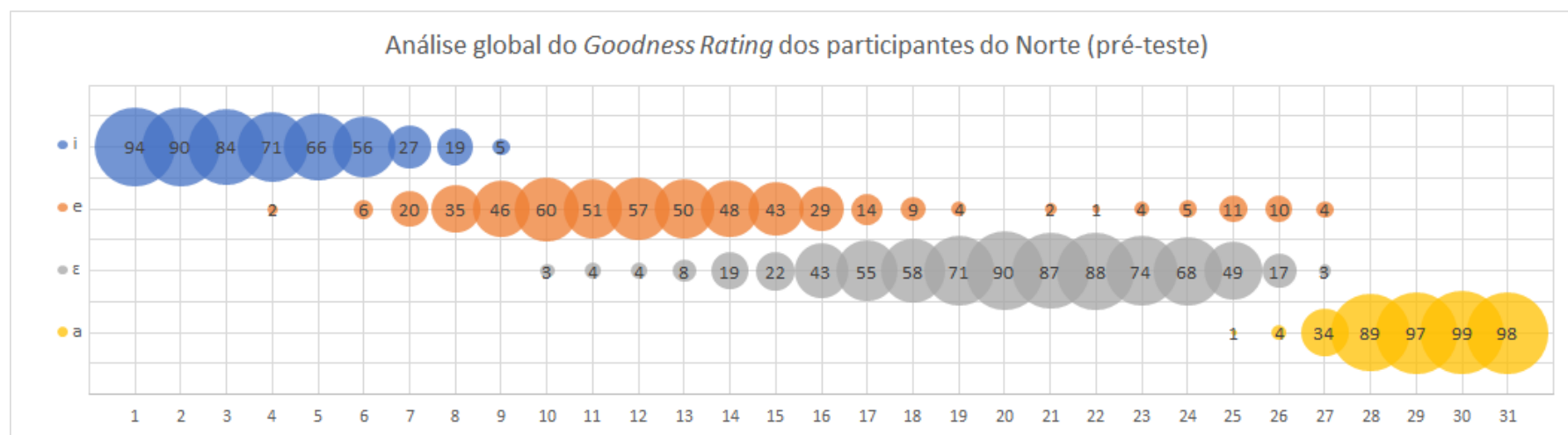


Figura 23: Representação do goodness rating mediano dos 6 informantes do Norte de Portugal.

Relativamente ao segundo fator, escolhemos metodologias experimentais nas quais se inclui o ponto de vista da teoria do protótipo, isto é, a classificação de *goodness*, enquanto que em Silva e Neves (2016) foram empregues metodologias baseadas na perceção categorial.

⁶⁰ Tendo apenas 6 amostras do Norte, e havendo também uma limitação temporal do presente estudo, não se realizou uma análise estatística dos dados recolhidos com os participantes nortenhos.

No que diz respeito à criação dos estímulos, em Silva e Neves (2016) foi utilizado o sintetizador KLSYN88 (Klatt & Klatt, 1990), manipulando os valores dos seguintes parâmetros: f_0 , F_1 , F_2 , F_3 e a duração. Por outras palavras, os seus estímulos não possuem informações acústicas além destes parâmetros marcados. Por sua vez, os estímulos desta investigação mantêm-se, tanto quanto possível, as informações acústicas da fala natural. É importante, no entanto, referir que o sintetizador utilizado neste trabalho, TANDEM-STRAIGHT (Kawahara *et al.*, 2009), ainda não nos permitiu controlar diretamente os valores dos formantes de todos os estímulos, como possibilita o sintetizador utilizado em Silva e Neves (2016).

6.4. Influência das relações fonológicas na percepção da fala: Suposição de Troubetzkoy (1969 [1939])

Refletimos, finalmente, se os resultados desta investigação sustentam a suposição de Troubetzkoy (1969 [1939]:78), isto é, se a percepção da fala é afetada não só pela presença e ausência dos sons no sistema de L1 do ouvinte, mas também pelas relações que os sons estabelecem no sistema. Afirmamos que, em princípio, os resultados sustentam positivamente a suposição. A fronteira 2 (/e/-/ε/), a oposição subjacente que se neutraliza no nível de superfície, demonstrou uma menor nitidez do que a fronteira 3 (/ε/-/a/), a oposição que nunca se neutraliza. A fronteira 1 (/i/-/e/) também exibiu uma menor nitidez do que a fronteira 3 na percepção dos falantes jovens de Lisboa, tendo sido levantada, com os resultados estatísticos, uma possibilidade de ser a menos nítida das três. Considerando a neutralização das oposições da abertura na posição inicial absoluta, por exemplo “[ε]/[i]colo’gista” (cf. Veloso, 2016: 639), este fenómeno perceptual da fronteira 1 pode ser explicado a partir da influência da relação fonológica. No entanto, esta hipótese é derivada de uma mera especulação, podendo haver várias outras hipóteses não fonológicas. É interessante considerar o fator particular deste dialeto. Assim, esta questão relativa à fronteira 1 fica em aberto.

6.5. *Perceptual hyperspace* (Johnson, 2000) e *adaptive dispersion* (Liljencrants & Lindblom, 1972)

Dois participantes da nossa investigação, os sujeitos BM e AC, quase não categorizaram, nem classificaram com uma pontuação alta, a categoria /i/ (ver as Figuras 18 e 19). Podíamos ter considerado estes dados *outliers*, mas os *goodness score* destes sujeitos revelam uma boa estruturação das outras categorias, com os protótipos no centroide. Assim sendo, estes sujeitos parecem preferir os sons da categoria /i/ mais periféricos, isto é, a categoria /i/ que não está incluída no *continuum* dos nossos estímulos.

Johnson *et al.* (1993) e Johnson (2000) observaram uma tendência dos ouvintes, numa tarefa de percepção da fala, para escolher os sons vocálicos periféricos, isto é, os estímulos com valores mais extremos do que os estímulos com os valores mais próximos das vogais produzidas por esses ouvintes. O mesmo fenómeno foi confirmado em Lively e Pisoni (1997), uma replicação da experiência de Kuhl (1992). Nesta experiência, na tarefa de *goodness ratings*, os participantes não classificaram como o melhor exemplo da categoria /i/ o estímulo com o valor médio da produção da voz masculina (que é suposto ser o protótipo desta categoria) mas os estímulos com o valor F2 mais alto do que o suposto protótipo. Johnson (2000) sugeriu duas alternativas para explicar esta tendência, o chamado efeito de *perceptual hyperspace*. A primeira baseia-se na proposta de Lindblom (1990): “the hyperspace effect reflects listeners’ production targets that are subject to undershoot in production” (Johnson, 2000: 182), isto é, uma falha em alcançar o objetivo pretendido, da produção. A segunda, por sua vez, considera esta tendência dos ouvintes como uma evidência, ainda que indireta, da hipótese de *adaptive dispersion* (Liljencrants & Lindblom, 1972): “(t)he hypothesis that the distinctive sounds of a language tend to be positioned in phonetic space so as to maximize perceptual contrast” (Johnson, 2000: 181). É necessário realizar mais estudos para definir a razão pela qual os dois sujeitos do nosso estudo (não) escolheram a categoria /i/.⁶¹

⁶¹ Para este problema do protótipo /i/ ver também Boersma (2005) em que se propõe um modelo de “fonologia e fonética em paralelo” baseado na *Optimality Theory* (Prince & Smolensky 1993).

Capítulo 7. – Conclusão e considerações finais

Neste último capítulo apresentam-se as conclusões, reflexões e críticas da investigação realizada e propõem-se algumas sugestões para investigações futuras.

7.1. Conclusões gerais

Partindo de dois pressupostos teóricos, o objetivo geral da investigação do presente trabalho era estudar a influência das relações fonológicas de L1 na perceção dos falantes nativos com a metodologia que engloba o ponto de vista da teoria do protótipo. Para este objetivo, as quatro vogais não recuadas, /i, e, ε, a/ do PE foram escolhidas como objeto da investigação, sendo as vogais médias anteriores consideradas vogais-alvo e as outras duas vogais, por sua vez, foram empregues como objeto de comparação. Tendo em consideração a relação complexa que estas vogais-alvo estabelecem entre as formas subjacentes e as de superfície, formulamos, então, a hipótese de que o grau de nitidez da fronteira entre as vogais médias anteriores é menor do que noutras fronteiras vocálicas na perceção dos falantes nativos de PE, mais especificamente, dos falantes jovens nativos de Lisboa. Hipótese esta que é sustentada por estudos anteriores relativos às vogais do PE e também pela teoria fonológica desta língua. De modo a testar a hipótese:

- 1) Analisaram-se as categorias distintas;
- 2) Analisou-se a estrutura interna, confirmando a existência do protótipo de cada categoria;
- 3) Analisou-se o grau da confusão na decisão da categorização nas fronteiras entre categorias.

Através das respostas obtidas na tarefa de identificação, concluiu-se que os participantes exibiram quatro categorias vocálicas distintas, que estão em conformidade com a teoria fonológica e com a observação num estudo acústico da produção desta língua. Verificaram-se, no entanto, variações individuais relativas às fronteiras entre categorias. A partir das respostas obtidas na tarefa de classificação de *goodness*, por sua vez, confirmou-se a estrutura interna de cada categoria, que possui o seu protótipo no centroide⁶² (em termos de número dos estímulos), variando também individualmente. Observando a variação individual,

⁶² Uma vez que os nossos estímulos da categoria /i/ e /a/ não abrangem além do valor produzido naturalmente, não conseguimos confirmar as suas categorias integrais. Aborda-se esta questão na secção 7.2. *Limitações da investigação* deste Capítulo.

quer nos resultados das fronteiras, quer nos resultados dos protótipos, parece haver uma afinidade entre a produção e a percepção. No entanto, não concluímos a ligação estreita estes dois domínios.

Para a nossa questão principal, isto é, se há influências das relações fonológicas na percepção dos falantes nativos de PE, obtiveram-se resultados positivos. O grau de nitidez das fronteiras entre /i/ e /e/, e /e/ e /ɛ/ revelaram uma menor nitidez na percepção dos participantes, em comparação com a fronteira entre /ɛ/ e /a/, sendo a última oposição não neutralizável. Podemos concluir que uma menor nitidez da fronteira /e/ e /ɛ/ se deve à influência das relações fonológicas complexas que estes sons estabelecem no sistema vocálico do PE. No que concerne à comparação entre a fronteira /i/ e /e/ com a fronteira /e/ e /ɛ/, os resultados estatísticos mostraram que a primeira pode ser menos nítida do que a segunda. Relativamente a um menor grau de nitidez da fronteira /i/ e /e/ em relação ao da fronteira /e/ e /ɛ/, pode resultar de uma especificidade do PE, ou de particularidades da metodologia utilizada no presente estudo. Existe ainda uma possibilidade de especificidade dialetal. De acordo com os comentários dos participantes lisboetas, e fazendo também uma breve comparação com o resultado dos nortenhos no pré-teste, este resultado sugere uma hipótese da influência do fator dialetal na percepção, o que é extremamente interessante. A questão fica em aberto para trabalhos futuros.

7.2. Limitações da investigação

Convém referir que o presente estudo teve as suas limitações, nomeadamente, o número das amostras e as características dos estímulos. Relativamente à primeira limitação, o facto do estudo se ter cingido apenas a um grupo pequeno (n=15) pode ter tido consequência nos resultados das análises estatísticas. No que diz respeito aos estímulos, apontamos dois pontos cruciais que podem ter induzido resultados desadequados às questões que queríamos esclarecer e/ou podem não nos ter permitido aprofundar questões de maneira mais rigorosa. Primeiro, a falta de alcance dos estímulos além de /i/ (estímulo 01) e de /a/ (estímulo 31) que completa a suposta categoria inteira destas duas vogais. Se o tivéssemos feito, poderiam realizar-se testes de comparação de variância que nos permitiriam analisar a amplitude e a variabilidade das categorias. Deste modo, oferecer-nos-iam evidências adicionais de forma a obter conclusões acerca da nitidez das fronteiras. Além disso, o alargamento do alcance da categoria /i/ possibilitaria a confirmação dos desempenhos dos dois sujeitos que quase não

identificaram a categoria /i/ com os estímulos utilizados no nosso estudo. Segundo, a distância acústica entre os estímulos foi também um fator que nos limitou na realização dos testes supramencionados. Tendo em conta a melhor percepção da fala natural em comparação com a sintetizada, foi escolhido, no nosso estudo, o *software* TANDEM-STRAIGHT (Kawahara *et al.*, 2009), que consegue gerar estímulos com as informações ricas da fala natural. Por outro lado, este ainda não nos permite manipular os valores dos formantes com precisão. Os estímulos desta investigação foram gerados a partir de duas âncoras da fala natural, criando 11 estímulos fisicamente mais ou menos equivalentes entre estas âncoras. No entanto, cada intervalo possui naturalmente uma distância física diferente: F₁ entre /i/ -/e/ (estímulo 01- 11) 57 Hz; /e/ - /ε/ (estímulo 11-21) 64 Hz e /ε/- /a/ (estímulo 21-31) 269 Hz⁶³. Deste modo, os nossos estímulos pertencentes ao último intervalo apresentam uma distância de F₁ 4 vezes maior do que os outros. Por este motivo, não foi possível realizar testes de comparação de variância, nem uma inspeção visual dos gráficos relativa ao número dos estímulos sobrepostos⁶⁴. Destacou-se mais o valor de F₁, pois esse deverá ter um papel importante na distinção das vogais tratadas neste estudo (embora esse facto não tenha sido ainda determinado), tendo em conta a relação observada entre os valores acústicos (F₁) na produção e a altura das vogais fonológicas (Escudero *et al.*, 2009: 1385)⁶⁵. Por outro lado, considerando a estrutura da experiência, sobretudo a duração, gerar 4 vezes mais estímulos para o último intervalo não seria a melhor solução. Uma vez que, com cerca de 60 estímulos, a duração da experiência tornar-se-ia muito maior, o que causaria dificuldades na concentração dos participantes ou até mesmo no seu recrutamento. Sugerimos, por exemplo, dividir em números diferentes para cada intervalo na realização de *morphing* com o TANDEM-STRAIGHT. Esperamos, por outro lado, o desenvolvimento tecnológico deste *software*, e/ou de outros sintetizadores que consigam gerar estímulos com alta qualidade e precisão, de modo a resolver o dilema da “faca de dois gumes”.

Sendo que a investigação trata da percepção auditiva, deve também ser considerado o uso de uma escala percetiva, como mel ou Bark, escalas que representam distâncias percetivas.

⁶³ Os sons referidos e os valores do intervalo são os que funcionaram como âncoras na criação dos estímulos e os valores da distância entre estes.

⁶⁴ Fizemos esta análise comparando o resultado da experiência principal e do pré-teste relativamente às transições das categorias de /i/ a /e/ e de /e/ a /ε/. Esta comparação é válida, dado que as distâncias comparadas são as mesmas.

⁶⁵ Relativamente aos valores de F₂ de cada intervalo dos nossos estímulos são menos regulares: entre /i/ -/e/ (estímulo 01- 11) 327 Hz; /e/ - /ε/ (estímulo 11-21) 58 Hz e /ε/- /a/ (estímulo 21-31) 599 Hz.

Acrescenta-se ainda que, na análise do grau de nitidez das fronteiras com os *goodness scores*, não foi possível diferenciar matematicamente o seguinte contraste: $50/50=1$ vs. $20/20=1$, mesmo que tenhamos uma certa intuição de que no primeiro caso haja maior confusão na decisão da categorização do que no segundo caso. E, por este motivo também, a realização de outros testes como os de comparação de variância será pertinente.

Todas as limitações metodológicas encontradas no nosso estudo, serão consideradas como indicações para melhorar trabalhos futuros.

7.3. Trabalhos Futuros

Tendo em conta o estudo que foi feito, salientam-se alguns aspetos merecedores de um trabalho futuro. Assim, o mesmo estudo pode ser alargado à variante do PB e a outros dialetos do PE, o que permite estabelecer comparações e constatar as suas especificidades na perceção, assim como proporcionar evidências ao modelo teórico cognitivo⁶⁶ e/ou fonológico.

Relativamente às especificidades dialetais do PE, o estudo de Rodrigues e Martins (1999) sobre o espaço acústico das vogais acentuadas dos falantes de Braga revelou um triângulo vocálico mais alargado, em comparação com o do dialeto de Lisboa⁶⁷, confirmando também que as vogais estão razoavelmente definidas em zonas acústicas distintas. Esta diferença acústica entre dialetos poderá ter consequências na categorização dos falantes de cada dialeto. Caso se avance com estudos comparativos entre estes dialetos e/ou outros do PE, seria igualmente pertinente obter, com a mesma metodologia, os dados acústicos da produção de cada dialeto⁶⁸ e também individuais. Desta forma, seria possível comparar diretamente a relação dos dois domínios: a perceção e a produção.

⁶⁶ Por exemplo, Hume e Johnson (2001) propuseram um modelo cognitivo global dos sons da fala, partindo do pressuposto relativo à interação entre o nível fonológico e a perceção- cf. secção 2.2. *Influência do sistema fonológico na perceção da fala*. De modo a verificar estas relações, os autores colocaram as três seguintes questões pertinentes: 1) até que ponto a perceção da fala influencia o sistema linguístico; 2) até que ponto o sistema linguístico influencia a perceção da fala; 3) onde é que os fenómenos da perceção da fala pertencem em relação a uma descrição formal da estrutura da língua (Hume & Johnson, 2001: 6).

⁶⁷ Rodrigues e Martins (1999) referiram os estudos de Delgado-Martins (1973) e de Andrade (1987, 1992) para definir as propriedades acústicas do vocalismo do português de Lisboa. (cf.-1999: 314)

⁶⁸ Relativamente aos dados acústicos da produção das vogais do português da variante brasileira, existe a investigação de Escudero *et al.* (2009) com participantes de São Paulo.

Seria interessante realizar-se um estudo experimental com as vogais médias recuadas que se neutralizam no contexto não acentuado e comparar o grau da nitidez das fronteiras com o das vogais médias anteriores⁶⁹.

As características dos estímulos gerados por diferentes sintetizadores e/ou fala natural podem ser investigadas. Por exemplo, poder-se-á investigar o mesmo fenómeno tratado no nosso estudo com estímulos gerados com um sintetizador.

Para terminar, reforçamos que as metodologias desta investigação, isto é, a avaliação do grau de nitidez da fronteira entre categorias com a classificação de *goodness* (*goodness scores*) foram, de certo modo, algo inovador e desafiante. É preciso ainda encontrar uma maneira de poder calcular matematicamente as diferenças do grau de nitidez das fronteiras.

⁶⁹ Tal como referido na secção 3.1. *As relações fonológicas e a percepção da fala*, a percepção das vogais médias recuadas foi investigada por Silva e Neves (2009).

Referências bibliográficas

Amorim, C. (2014). *Padrão de Aquisição de Contrastos do PE: a interação entre traços, segmentos e sílabas* (Dissertação de doutoramento). Porto: Faculdade de Letras da Universidade do Porto.

Acedida em <https://repositorium.sdum.uminho.pt/bitstream/1822/32998/3/Yang%20Shu.pdf>

Aaltonen, O., Eerola, O., Hellström, Å., Uusipaikka, E., & Lang, H. A. (1997). Perceptual magnet effect in the light of behavioral and psychophysiological data. In *The Journal of the Acoustical Society of America*. 101, 1090-1103.

Alvarado, N., & Jameson, K. (2002). The Use of Modifying Terms in the Naming and Categorization of Color Appearances in Vietnamese and English. In *Journal of Cognition and Culture*. 2(1). 53–80.doi:10.1163/156853702753693307

Barros, R.Q. (1994). *Contributo para a Análise Sociolinguística do Português de Lisboa: Variantes de /e/ e de /ɛ/ em Contexto pré-palatal* (Dissertação de Mestrado). Lisboa: FLUL.

Berlin, B., & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley & Los Angeles: University of California Press.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. In *Journal of Phonetics*, 20 (3), 305-330.

Best, C., & Tyler, M. (2007). Nonnative and Second Language Speech Perception: Commonalities and Complementarities. In O. Bohn, & M. Munro (Eds), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*. Amsterdam/Philadelphia: John Benjamins Publishing Company. pp. 13-34

Bisol, L. (2001). *Introdução a estudos de fonologia do português brasileiro*. (3ª ed.). Porto Alegre: Edipucrs.

Bisol, L. & Veloso, J. (2016). Phonological Processes Affecting Vowels: Neutralization, Harmony, and Nasalization. In: W. Leo Wetzels, Sergio Menuzzi, João Costa (Eds.). *The Handbook of Portuguese Linguistics*. Malden MA/Oxford UK: Wiley/Blackwell. pp. 69-84

Boersma, P. (2015). Prototypicality judgments as inverted perception. In G. Fanselow, F. Caroline, M. Schlesewsky, & R. Vogel. (Eds.). *Gradience in Grammar: Generative Perspectives*. New York: Oxford University Press. Retirado de <http://www.fon.hum.uva.nl/paul/papers/Prototypicality.pdf>

Bond, Z. S. (1999). *Slips of the ear: Errors in the perception of casual conversation*. San Diego, CA, US: Academic Press.

Bonilha, G. F. G. (2004). Sobre a aquisição das vogais. In Regina R. LAMPRECHT. (Ed.). *Aquisição fonológica do português: Perfil de desenvolvimento e subsídios para a terapia*. Porto Alegre: ARTMED. 61-71.

Boomershine, A., Hall, K. C., Hume, E., & Johnson, K. (2008). The impact of allophony vs. contrast on speech perception. In P. Avery, E. Dresher, & K. Rice (Eds.), *Phonological contrast: Perception and acquisition*. New York: Mouton de Gruyter. pp. 146–172

Boomershine, A. (2013). The perception of English vowels by monolingual, bilingual, and heritage speakers of Spanish and English. In Chad Howe, Sarah E. Blackwell, & Margaret Quesada (Eds.), *Selected Proceedings of the 15th Hispanic Linguistics Symposium*. Massachusetts: Cascadilla Press. pp 103 - 118.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. In *The Journal of the Acoustical Society of America*. 101, 2299-2310.

Brandão de Carvalho, J. (2011). Contrastive hierarchies, privative features, and portuguese vowels. In *Linguística: Revista de Estudos Linguísticos da Universidade do Porto*. 6:1. Porto: Faculdade de Letras da Universidade do Porto. 51-66

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. In *The Journal of Abnormal and Social Psychology*, 49(3), 454-462.

Câmara Jr., J. M. (1970). *Estrutura da Língua Portuguesa*. (29ª ed.) Petrópolis, RJ: Editora Vozes.

Castro, S.F.L (1993) Perceção categorial num continuum vozeamento em fala natural portuguesa, In *Actas do 1º encontro de processamento de língua portuguesa escrita e falada*, 15-18

Chomsky, N.& Halle, M. (1968). *The Sound Pattern of English*. New York NY: Harper & Row.

Clements, G. N. (1991) Vowel height assimilation in Bantu languages. In: *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: Special Session on African Language Structures*. Berkeley: University of California. 25-64

Cutler, A. (2012) *Native Listening Language: Experience and the Recognition of Spoken Words*. Cambridge, Mass.: MIT Press.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. In *Journal of the Acoustical Society of America*. 116(6). 3668-3678.

Denes, P.B. & Pinson, E.N. (1993). *The Speech Chain: The Physics and Biology of Spoken Language*. New York, N.Y: W.H. Freeman.

Dresher, B. E. (2004). *On the acquisition of phonological contrasts*. In J. van Kampen & S. Baauw (eds.) *Proceedings of GALA 2003*, Volume 1, Utrecht, LOT, 27–46.

Dresher, B. E. (2003). *The Contrastive Hierarchy in Phonology*. In *Tronto working in Linguistic (Special issue on Contrast in Phonology)*. Tronto: University of Tronto. 20. 47-62.

Eorola, O., Savela, J., Laaksonen, J.P.& Aaltonen, O. (2012). The effect of duration on vowel categorization and perceptual prototypes in a quantity language In *Journal of Phonetics* 40. pp. 315-328

Eerola, O., Savela, J. & Aaltonen O. (2014). Weighted vowel prototypes in Finnish and German. In *Journal of the Acoustical Society of America*. 135 (3). 1530-40. doi: 10.1121/1.4864305. Disponível na internet

https://www.researchgate.net/publication/260643943_Weighted_vowel_prototypes_in_Finnish_and_German Acedido a: 15/05/2018

Escudero, P., Boersma, P., Rauber, A., & Bion, R. (2009). A cross-dialect acoustic description of vowels: Brazilian versus European Portuguese. In *Journal of the Acoustical Society of America*. 126. 1379–1393.

Evans, B.G. & Iverson, P. (2004). Vowel normalization for accent: A comparison of northern and southern British English speakers. In *The Journal of the Acoustical Society of America*. 115. 352-361

Evans, B.G. & Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *The Journal of the Acoustical Society of America*. 121. 3814-3826. <https://doi.org/10.1121/1.2722209>

Falé, I. (2005). Parte II. Fonética In *Fonética e Fonologia do Português*. Lisboa: Universidade Aberta. 41-154.

Fant, G. (1960) *Acoustic theory of speech production*. The Hague, The Netherlands: Mouton.

Fernández, E.M. & Carins, H.S. (2011). *Fundamentals of psycholinguistics*. Wiley-Blackwell.

Ferrari, L. (2011). Categorização. In *Introdução à Linguística Cognitiva*. São Paulo: Contexto. Pp. 31-48.

Fikkert, J.P. M. (2005). From phonetic categories to phonological features specification: Acquiring the European Portuguese vowel system. In *Lingue e Linguaggio*. vol. 4. iss. 2, (2005), pp. 263-280

Fikkert, J.P.M.& Freitas, M.J. (2006). Allophony and allomorphy cue phonological development: Evidence from the European Portuguese vowel system In *Catalan Journal of Linguistics*. vol. 5. pp. 83-108

Freitas, M.J. (2004). The Vowel [i] in the Acquisition of European Portuguese. In J. Van Kanpen and S. Baauw (eds.) *Proceedings of GALA 2003*, vol.1, Utrecht: LOT pp 163-174

Greene, B. G. (1986). Perception of synthetic speech by nonnative speakers of English. *Proceedings of the Human Factors Society*, Vol. 2, pp.1340–1343.

Grieser, D., & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. In *Developmental Psychology*. 25. 577-588.

Harris, J. (1994). *English Sound Structure*. Cambridge: Blackwell.

Harris, J. & Lindsey, G. (2000). Vowel patterns in mind and sound. In Noel Burton-Roberts, Philip Carr & Gerard DOCHERTY (eds.). *Phonological knowledge: Conceptual and empirical issues*. New York: Oxford University Press.

Hillenbrand, J. M. (2013). Static and Dynamic Approaches to Vowel Perception. In G. S. Morrison and P. F. Assmann (eds.), *Vowel Inherent Spectral Change, Modern Acoustics and Signal Processing*, Berlin Heidelberg: Springer. 9- 30. DOI: 10.1007/978-3-642-14209-3_2

Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. In *Attention, Perception & Psychophysics*. 72(5), 1218–1227. <http://doi.org/10.3758/APP.72.5.1218>

Huang, T. (2001). The Interplay of Perception and Phonology in Tone 3 Sandhi in Chinese Putonghua. In *Linguistics*. 55. OSUWPL. 23-42

Hume, E., & Johnson, K. (2001). A Model of the Interplay of Speech Perception and Phonology. In *The Role of Speech Perception in Phonology*. The Netherland: Leiden. pp. 3-26

Hume, E.& Johnson, K. (2003). The impact of partial phonological contrast on speech perception. In *Proceedings of the Fifteenth International Congress of Phonetic Sciences. Barcelona 3-9 Augusto 2003*. Australia: Causal Productions. pp. 2385-2388.

Iverson, P. & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. In *The Journal of the Acoustical Society of America* 97(1):553-62

Iverson, P. & Kuhl, P. K. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. In *The Journal of the Acoustical Society of America* 99 (2):1130-40

Iverson, P. & Kuhl, P. K. (2000). Perceptual magnet and phoneme boundary effects in speech perception: Do they arise from a common mechanism? In *Perception & Psychophysics*. 62: 874

Iverson, P., Wagner, A. & Rosen, S. (2016). Effects of language experience on pre-categorical perception: Distinguishing general from specialized processes in speech perception In *The Journal of the Acoustical Society of America*. 139. 1799- 1809
<https://doi.org/10.1121/1.4944755>

Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates*. Cambridge, Mass.: MIT Press.

Johnson, K. (2010). *Acoustic and Auditory Phonetics*. (3^a ed) - Chapter 5. Hoboken, NJ: Wiley-Blackwell. Retirado de <https://escholarship.org/uc/item/0p18z2s3>

Johnson, K. (2000). Adaptive dispersion in vowel perception. In *Phonetica* 57:181-188

Johnson, K., Flemming, E. & Wright, R. (1993). The hyperspace effect: phonetic targets are hyperarticulated. In *Language*. 69. 505–528.

Kawahara, H; Takahashi, T; Morise, M & Banno, H (2009). Development of exploratory research tools based on TANDEM-STRAIGHT In *Proceedings: APSIPA ASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference*. Japão: Universidade de Hokkaido.111-120.

Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. In *Percept Psychophys*. 50(2), 93-107

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. In *Nature Reviews Neuroscience*. 5. 831-843.

Kuhl, P.K., Conboy, B.T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early Speech Perception and Later Language Development: Implications for the “Critical Period”. In *Language Learning and Development*. 1(3&4), 237–264

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). In *Philosophical Transactions of the Royal Society B*, 363, 979-1000.

Kuhl, P. K. & Iverson (1995). Linguistic experience and the Perceptual Magnet Effect In W. Strange (ed). *Speech Perception and Linguistic Experience*. pp 121-154

Kuhl, P.K. & Miller, J.D. (1975) Speech Perception by the Chinchilla: Voiced-Voiceless Distinction in Alveolar Plosive Consonants. In *Science, New Series*, Vol. 190, No. 4209. pp. 69-72.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. In *Science* 255, 606–608

Labov. W. (1973). The boundaries of words and their meanings. In C.-J. Bailey and R. Shuy (eds.), *New Ways of Analyzing Variation in English*. Washington, DC: Georgetown U. Press. Pp. 340-373. Reprinted in Aarts. B. et al. (eds.), *Fuzzy Grammar*. Oxford: Oxford U. Press. Pp. 67-90

Ladefoged, P. (2005). *Vowels and consonants: An introduction to the sounds of languages*. Malden, MA: Blackwell Pub.

Ladefoged, P., & Johnson, K. (2011). *A course in phonetics*. (6^a ed.). Boston, MA: Wadsworth.

Lee, S. H. (2010). Contraste das vogais no PB e OT. In *Estudos Linguísticos*. 39(1) São Paulo. 35-44

Lieberman, A., Harris, K., Hoffman, H., Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. In *Journal of Experimental Psychology: Human Perception and Performance*, v.58, p. 358-368.

Liljencrants, J.; Lindblom, B. (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. In *Language* 48: 839–862

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In Hardcastle, Marchal. *Speech production and speech modeling*. Dordrecht: Kluwer Academic Publishers. pp. 403–439

Liu, C., & Kewley-Port, D. (2004). Vowel formant discrimination for high-fidelity speech In *The Journal of the Acoustical Society of America*. 116(2), 1224-33

Lively, S. E. & Pisoni, D. B. (1997). On Prototypes and Phonetic Categories: A Critical Assessment of the Perceptual Magnet Effect in Speech Perception. In *Journal of Experimental Psychology. Human Perception and Performance*. 23(6), 1665–1679.

Logan, J. S., Greene, B. G., & Pisoni, D. B. (1989). Segmental intelligibility of synthetic speech produced by rule. In *The Journal of the Acoustical Society of America*. 86. 566–581.

MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. In *Applied Psycholinguistics*. 2(4), 369-390.

Maddieson, I. (1984) *Patterns of Sounds*. New York: Cambridge University Press.

Mateus, M. H. & Andrade, E. (2000). *The Phonology of Portuguese*. Oxford: Oxford University Press.

Mateus, M. H. M. (2003). Aspectos Fonológicos e Prosódicos da Gramática do Português, parte VI. 25. In *Gramática da Língua Portuguesa*. (5ª ed). Lisboa: Caminho. pp. 987- 1053

Masapollo, M., Polka, L. & Molnar, M. (2017). Directional asymmetries reveal a universal bias in adult vowel perception. In *The Journal of the Acoustical Society of America*. 141(4):2857-2869. DOI: 10.1121/1.4981006

McAuliffe, M. (2016). *STRAIGHT tutorial*.

https://memcauliffe.com/straight_workshop/index.html consultado em 20 de setembro de 2018.

Miller, J.L. (1994). On the internal structure of phonetic categories: a progress report. *Cognition*, ISSN: 0010-0277, Vol: 50, Issue: 1, Page: 271-285

Miller, J. L., & Volaitis, L. E. (1989). Effect of speaking rate on the perceptual structure of a phonetic category In *Perception & Psychophysics*. 46. 505-512.

Miyawaki, K, Jenkins, J., Strange, W., Verbrugge, R.R., Liberman, A.M.J., Fujimura, O. (1975). An effect of linguistic experience: the discrimination of [r] and [l] by native speakers of Japanese and English In *Perception and Psychophysics*. 18 (5) 331–340.

Miyaoka, O. (2006). 言語の違い 認識の違い In 宮岡伯人編 今、世界のことばが危ない! : グローバル化と少数者の言語: 2005 第19回「大学と科学」公開シンポジウム講演収録集. 東京 : クバプロ (As diferenças entre as línguas, as diferenças de reconhecimento. In Miyaoka, O. (Ed.) *As línguas do mundo estão em risco!: Coleção das palestras do 19º simpósio em 2005 “As Universidades e Ciências”*. Tokyo: Kubapro. Tradução do título pela autora da dissertação.) pp. 43-49

Mochizuki, M. (1981). The identification of /r/ and /l/ in natural and synthesized speech. In *Journal of Phonetics*, 9(3), 283-303.

Newman, R.S. (2003). Using links between speech perception and speech production to evaluate different acoustic metrics: A preliminary report. In *The Journal of the Acoustical Society of America*. 113(5):2850-60.

Oden, G. C., & Massaro, D. W. (1978). Integration of featural information in speech perception. In *Psychological Review*. 85(3), 172-191.

<http://dx.doi.org/10.1037/0033-295X.85.3.172>

Oliveira, D. M. (2016). *Percepção e Produção de Sons Consonânticos do Português Europeu por Aprendentes Chineses* (Dissertação de mestrado). Braga: Universidade do Minho. Acedida em

https://repositorium.sdum.uminho.pt/bitstream/1822/41626/1/Disserta%C3%A7%C3%A3o_d_e_mestrado.pdf

Peperkamp, S., & Dupoux, E. (2000). Coping with phonological variation in early lexical acquisition. In Lasser (ed.). *The Process of Language Acquisition*. Berlin: Peter Lang Verlag, pp. 359-385.

Disponível na internet

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.16.6852&rep=rep1&type=pdf>

Acedido a: 01/07/2018.

Phụng, B. (1998). *English- Vietnamese Dictionary*. Nhà xuất bản văn hóa

Pisoni, D.B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels In *Perception & Psychophysics* 13: 253-260

Pisoni D.B. (1997). Perception of Synthetic Speech. In: van Santen J.P.H., Olive J.P., Sproat R.W., Hirschberg J. (eds) *Progress in Speech Synthesis*. New York, NY: Springer.

Pisoni, D. B., Manous, L. M., and Dedina, M. J. (1987). Comprehension of natural and synthetic speech: Effects of predictability on the verification of sentences controlled for intelligibility. In *Computer Speech & Language*. Vol 2. 303–330.

Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. In *Journal of the Acoustical Society of America*. 89(6). 2961-2977.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. In *Journal of Experimental Psychology: Human Perception and Performance*. 20(2), 421-435.

Prince, A. & Smolensky, P. (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*.

Ralston, J. V., Pisoni, D. B., Lively, S. E., Greene, B. G., and Mullennix, J. W. (1991). Comprehension of synthetic speech produced by rule: Word monitoring and sentence-by-sentence listening times. In *Human Factors* 33. 471–491.

Rauber, A. S. (2008). An acoustic description of Brazilian Portuguese oral vowels. In *Diacritica, Ciências Da Linguagem* 22(1), Braga: CHEUM. 229-238.

Repp, B. H., & Liberman, A. M. (1987). Phonetic category boundaries are flexible. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition*. New York, NY, US: Cambridge University Press. pp. 89-112

Rosch, E. H. (1973). Natural categories. In *Cognitive Psychology*. 4(3). 328-350.

Rosch, E. H. (1975). Cognitive reference points. In *Cognitive Psychology*. 7(4). 532-547

Rosch, E. (1978). Principles of categorization. In: E. Rosch, & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum. pp. 28-49

Rodrigues, C. (2003). *Lisboa e Braga: Fonologia e Variação*. Lisboa: Fundação Calouste Gulbenkian.

Rodrigues, C. & Martins, F. (1999). Espaço acústico das vogais acentuadas de Braga. In *Actas do XV Encontro Nacional da Associação de Linguística*. vol. II. Faro: Universidade do Algarve. pp. 301-317

Rosner, B. S., & Pickering, J. B. (1994). *Oxford psychology series, No. 23. Vowel perception and production*. New York, NY, US: Oxford University Press.

<http://dx.doi.org/10.1093/acprof:oso/9780198521389.001.0001>

Samuel, A. G. (1982). Phonetic prototypes. In *Perception & Psychophysics*. 31. 307-314

Schwab, E. C., Nusbaum, H. C., & Pisoni, D. B. (1985). Some effects of training on the perception of synthetic speech. In *Human Factors*. 27(4), 395-408.

Silva, A.S. (1997). A linguística cognitiva - Uma breve introdução a um novo paradigma em linguística. In *Revista portuguesa de humanidades*. Vol. 1. Nº 1-2. 59-10

Silva, D.M.R. & Neves, R.R. (2009). Um estudo experimental sobre a percepção do contraste entre as vogais médias posteriores do português brasileiro In *DELTA*. vol.25: 2. pp.319-345

Silva, D.M.R. & Neves, R.R. (2016). Perception of height and categorization of Brazilian Portuguese front vowels In *DELTA* vol.32 nº2 May/ Aug, pp.355-373

Shu, Y. (2014). *Percepção das Consoantes Oclusivas de Português L2 sob a Influência de Mandarim L1* (Dissertação de mestrado). Braga: Universidade do Minho. Acedida em <https://repositorium.sdum.uminho.pt/bitstream/1822/32998/3/Yang%20Shu.pdf>

Stevens, K. N., Libermann, A. M., Studdert-Kennedy, M., & Ohman, S. E. (1969). Cross language study of vowel perception. In *Language and Speech*. 12(1), 1-23.

Tatham, M. & Morton, K (2006). *Speech Production and Perception*. UK: Palgrave Macmillan

Toda, T. (2003). Acquisition of Special Morae in Japanese as a Second Language (L2 Acquisition, <Feature Articles> Phonetics and Phonology of Language Acquisition) in *音声研究 (Onsei kenkyu)*, volume 7, 70-83

Trubetzkoy, N. S. (1939). *Grundzuge der Phonologie*. Trad. fr. de J. Cantineau. *Principes de phonologie*. Paris: Klincksieck, 1969.

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C.T. (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6-12 months: language-specific developmental changes in speech perception abilities. in *The 3rd International Conference on Spoken Language Processing*, September 18-22, 1994, Yokohama, Japan: ICSL.

Veloso, J. (2010). Central, epenthetic, unmarked vowels and schwas: A brief outline of some essential differences. In *Linguística – Revista de Estudos Linguísticos da Universidade do Porto*. v. 5. n. 1. Porto: Faculdade de Letras da Universidade do Porto.193-213

Veloso, J. (2012). Vogais centrais do português europeu contemporâneo: uma proposta de análise à luz da fonologia dos elementos In *Letras de Hoje*. 47:3. Porto Alegre: Escola de Humanidade: Pontifícia Universidade Católica do Rio Grande do Sul. 234- 243

Veloso, J. (2015) *Introducao a Fonologia. Nivel fonetico e nivel fonologico*. Porto: Faculdade de Letras da Universidade do Porto (ms.) (doi: 10.13140/RG.2.1.1591.8805).

Veloso, J. (2016) O sistema vocálico e a redução e neutralização das vogais átonas em português europeu contemporâneo. In: Ana Maria Martins, Ernestina Carrilho (Eds.). *Manual de Linguística Portuguesa*. Berlin: De Gruyter. pp. 636–662.

Werker, J. F. & Lalonde, C. (1988). Cross-language speech perception: initial capabilities and developmental change. In *Development Psychology* 24, 672–683

Werker, J. F. & Tees, R. C (1984a) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. In *Infant Behavior & Development* 25 (2002) 121–133

Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. In *The Journal of the Acoustical Society of America*. 75(6), 1866–1878.

Wetzels, L. (1992). Mid vowel neutralization in Brazilian Portuguese in *Cadernos de Estudos Lingüísticos*, Campinas. 23: pp. 19-55.

Wetzels, W. L. M. (2011) The representation of vowel height and vowel height neutralization in Brazilian Portuguese (Southern Dialects). In J. A. Goldsmith, E. Hume, & W. L. M. Wetzels (Eds.), *Tones and Features: Phonetic and Phonological Perspectives* pp. 331-360

Computer program:

Boersma, P., & weenink, D. (2017). “PRAAT: Doing phonetics by computer”. *Version 6.0.37*, <http://www.praat.org/>

Kawahara *et al.* (2009) *TANDEM- STRAIGHT*

Rstudio version 1.1.453, <https://www.rstudio.com/products/rstudio/download/>

Anexos

Anexo 1 – Questionário sobre o perfil biográfico e linguístico dos participantes

Background Questionnaire QUESTIONÁRIO

Nome:

_____ Grau

de Escolaridade: _____ Ano:

Telf.: _____ E-mail: _____ 1.

Idade: _____ 2. Sexo: M/ F

3. Naturalidade:

4. Residência (cidade): _____

4.1. Reside sempre nesta localidade? Sim/ Não

5. A sua língua materna é o português europeu? Sim/ Não

6. Ambos os pais são portugueses? Sim/ Não

7. Tem dificuldade/ problema auditivo? Sim/ Não

8. Domina outra língua além do português europeu? Sim/ Não

8.1. Se respondeu sim,

8.1.2. Qual língua?

8.1.3. Qual é o seu nível de proficiência da língua referida acima, segundo o QECRL?

9. Nunca teve o treino da fonética. Sim/ Não

Lisboa, _____

Anexo 2– Autorização

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Concordo em participar, como voluntário, no estudo que tem como pesquisador responsável, Megumi Im, aluna de Mestrado em Linguística da Faculdade de Letras da Universidade do Porto, que pode ser contactado pelo e-mail pame1967@hotmail.com e pelo telefone 917 117 660. Tenho em consideração que o estudo tem em vista realizar gravações, visando, por parte da referida aluna a realização de uma dissertação do mestrado supramencionado intitulada “Experiência de perceção das vogais semifechadas e semiabertas pelos falantes nativos de PE”. A minha participação consistirá em conceder gravações de fala natural e das quator vogais do PE. Entendo que o estudo a realizar possui como finalidade a pesquisa académica e atesto que os dados obtidos não serão divulgados, a não ser com prévia autorização, e que nesse caso será preservado o anonimato dos participantes, assim como a minha privacidade. Além disso, sei que posso abandonar minha participação na pesquisa quando quiser e que não receberei nenhum pagamento por esta participação.

Assinatura

Lisboa , ___ de _____ de 2017

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Concordo em participar, como voluntário, no estudo que tem, como investigador responsável, Megumi Im, aluna de Mestrado em Linguística da Faculdade de Letras da Universidade do Porto, que pode ser contactada pelo e-mail pame1967@hotmail.com e pelo telefone 917 117 660. Estou ciente de que farei um teste de perceção das vogais, visando, por parte da referida aluna, a realização de uma dissertação do mestrado supramencionado intitulada “Experiência de perceção das vogais semifechadas e semiabertas pelos falantes nativos de PE”. Entendo que o estudo a realizar possui como finalidade a pesquisa académica e que os dados daí resultantes serão analisados apenas para fins académicos, garantindo o meu anonimato. Além disso, sei que posso cessar a minha participação na pesquisa quando quiser e que não receberei nenhum pagamento por esta participação.

E por ser expressão da verdade, firmo a presente autorização que vai datada e assinada para que surta os devidos efeitos legais.

Lisboa/ Braga de 2018.

Nome do participante:

Assinatura do participante:

Anexo 3— Dados biográficos, linguísticos, sociodemográficos e comentários dos participantes no pré-teste

	Sexo	Idade	Natural	L1	Estatuto social	Grau académico	Língua Estrangeira	Feedback
AB	M	22		PE	Estudante universitário	Licenciatura(2º)	EN(B1), JP(A2)	
VB	F	21	Braga	PE	Estudante universitário	Licenciatura(3º)	EN(A2), JP(A1)	
FB	M	20	Braga	PE	Estudante universitário	Licenciatura(2º)	EN(C1)	Confusão entre as vogais [e] e [ɛ]
NB	F	24	Porto	PE	Investigador	Doutoramento (1º)	EN(C1), JP(A2)	
JB	M	24	Porto	PE	Estudante universitário	Mestrado (1º)	EN(C1)	
SB	F	23	Braga	PE	Estudante universitário	Mestrado (1º)	EN(C2)	

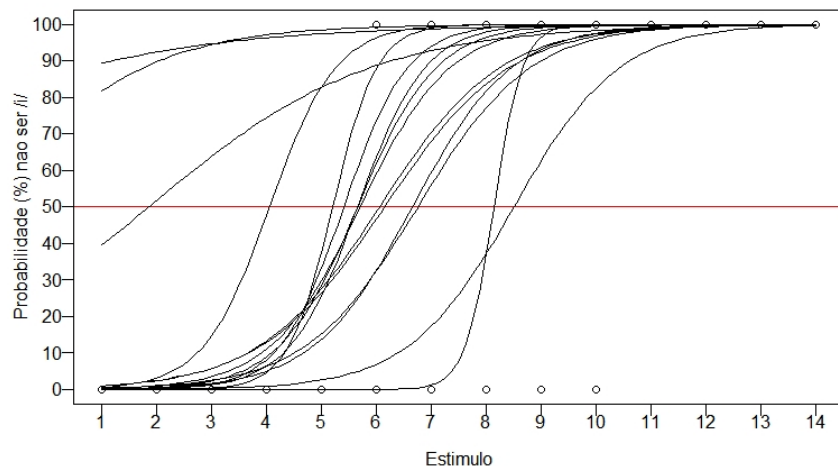
Anexo 4– Dados biográficos, linguísticos, sociodemográficos e comentários dos participantes na experiência principal

	Sexo	Idade	Natural	L1	Estatuto social	Grau académico	Língua Estrangeira	Feedback
CC	F	19	Lisboa	PE	Estudante universitário	Licenciatura(1º)		
AP	M	26	Lisboa	PE	Trabalhador	Licenciado em 2012		[i] e [e]foi difícil,sobretudo ao identificar [e] de "medo".
BM	F	20	Lisboa	PE	Estudante universitário	Licenciatura(3º)	EN(C2)	
MM	F	25	Lisboa	PE	Trabalhador	Mestre	EN (B1), SP(B1)	[i] e [e]foi difícil
MF	F	20	Lisboa	PE	Estudante universitário	Licenciatura(3º)	EN (A2)	[i] e [e]foi difícil.
ML	M	19	Lisboa	PE	Estudante universitário	Licenciatura(2º)	EN(B2)	
CP	F	34	Lisboa	PE	Investigador	Doutoramento	EN(C1)	
IM	F	20	Lisboa	PE	Estudante universitário	Licenciatura(2º)	EN(C2)	
MC	F	21	Lisboa	PE	Estudante universitário	Licenciatura(3º)	EN(C1)	
GS	M	28	Lisboa	PE	Estudante universitário	Licenciatura(2º)	EN(C2)	
FM	M	28	Lisboa	PE	Trabalhador	Mestre	EN(C1)	
AC	M	27	Lisboa	PE	Investigador	Mestrado (2º)	EN(C1), ES(B1)	

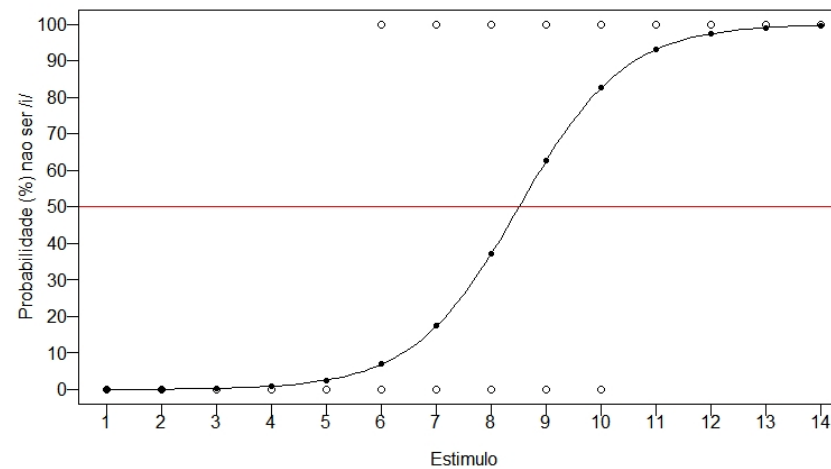
	Sexo	Idade	Natural	L1	Estatuto social	Grau académico	Língua Estrangeira	Feedback
BR	F	22	Lisboa	PE	Estudante universitário	Licenciatura(1º)	EN(C2)	
CL	F	19	Lisboa	PE	Estudante universitário	Licenciatura(2º)	EN(B2), FR(A1)	[i] foi difícil, porque normalmente não é tão longa.
LP	M	28	Lisboa	PE	Investigador	Mestrado(2º)	EN(C2)	

Anexo 5– Gráficos: Teste principal

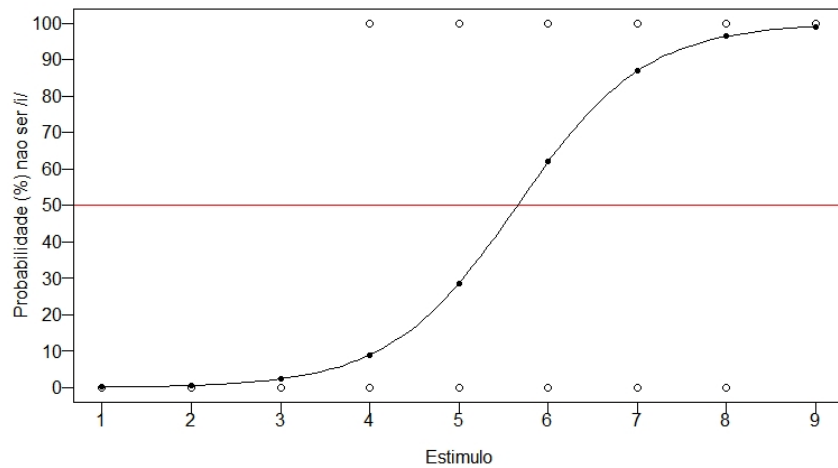
A curva de regressão logística para a transição da categoria /i/ para não /i/ (1/4)



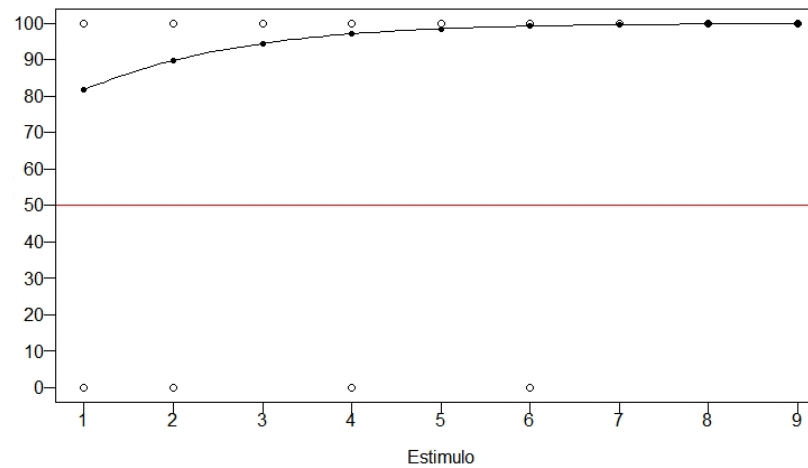
Todos os participantes



Sujeito CC

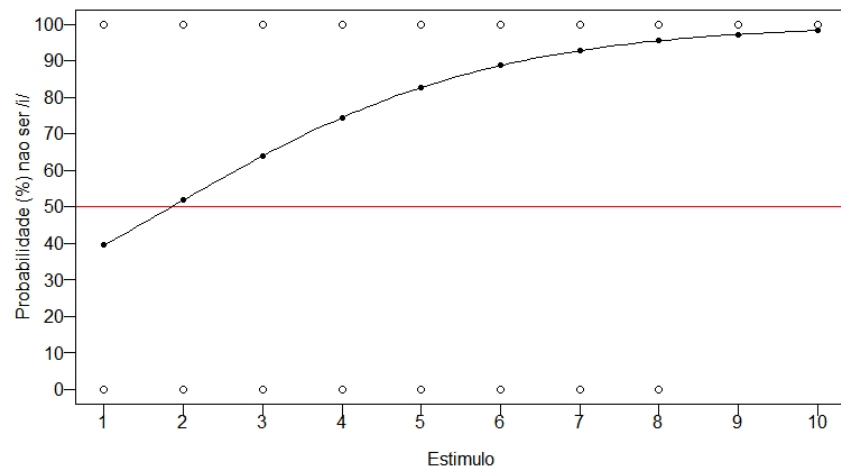


Sujeito AP

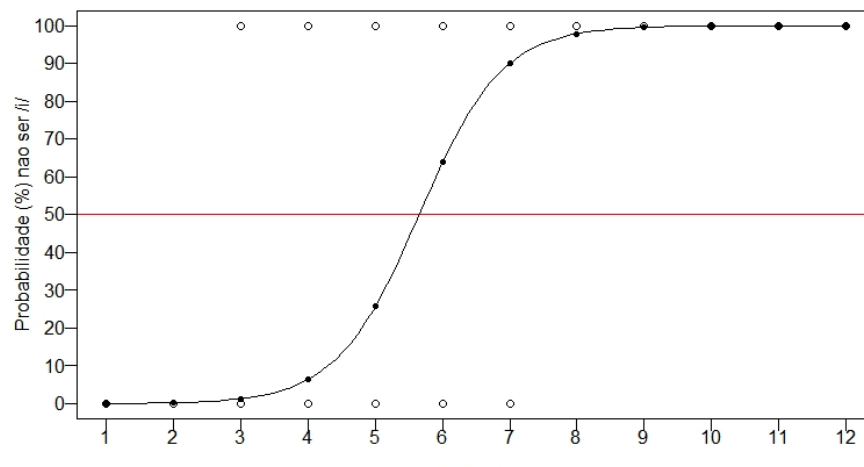


Sujeito BM

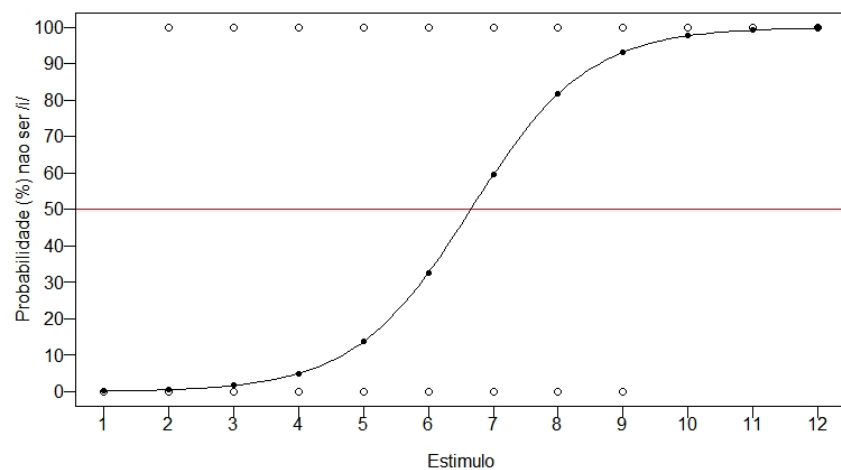
A curva de regressão logística para a transição da categoria /i/ para não /i/ (2/4)



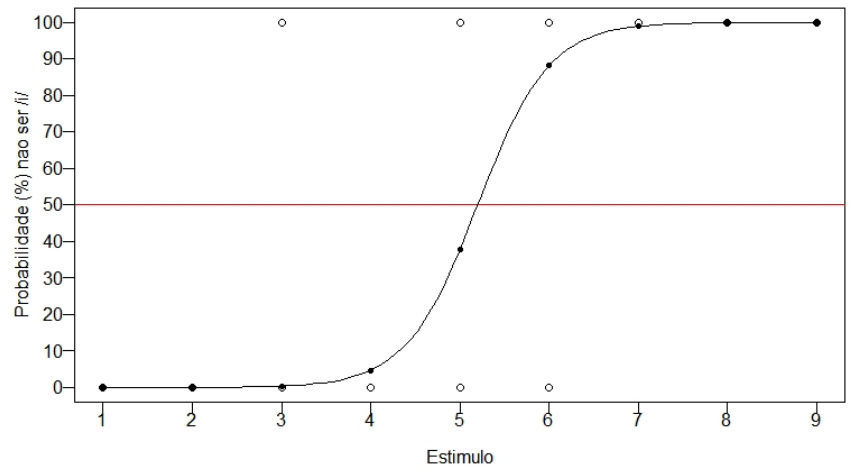
Sujeito MM



Sujeito MF

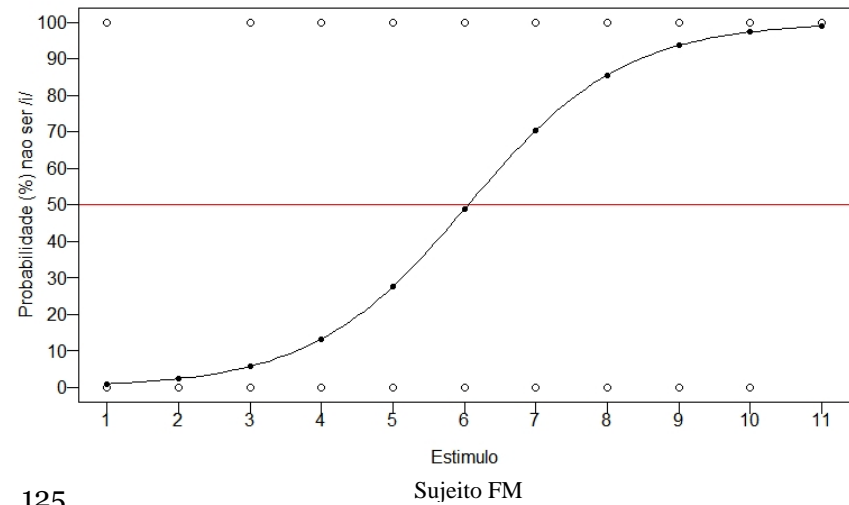
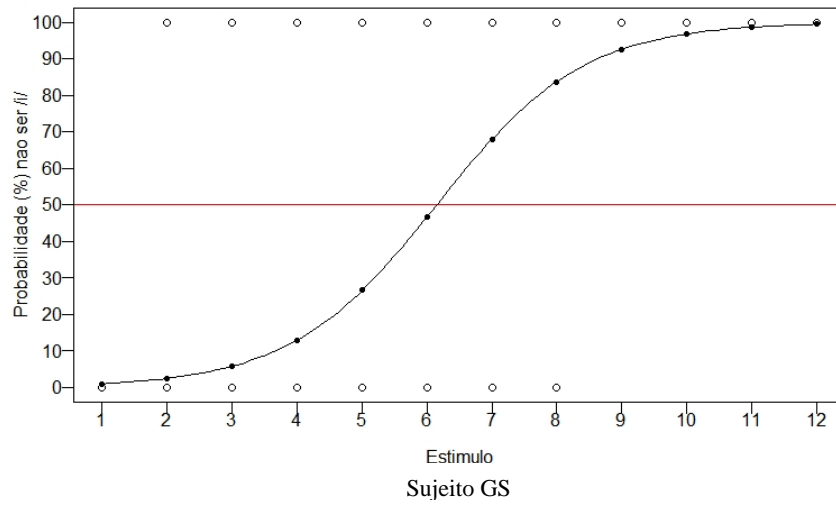
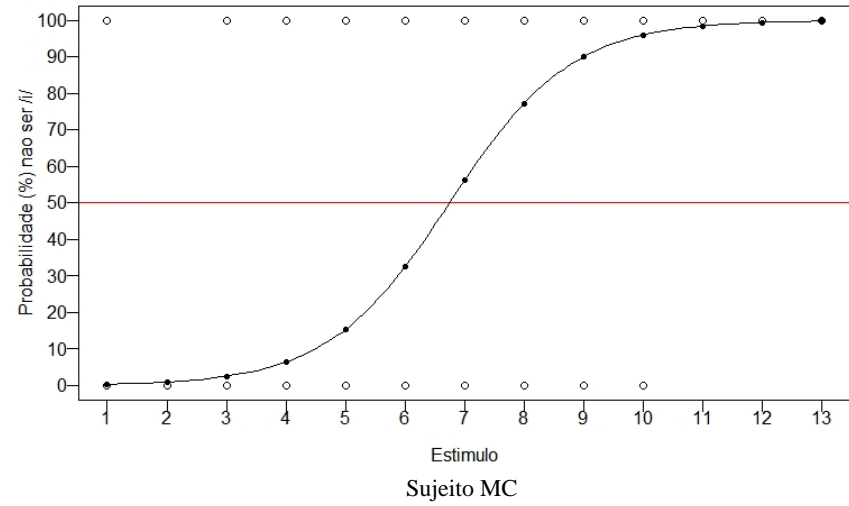
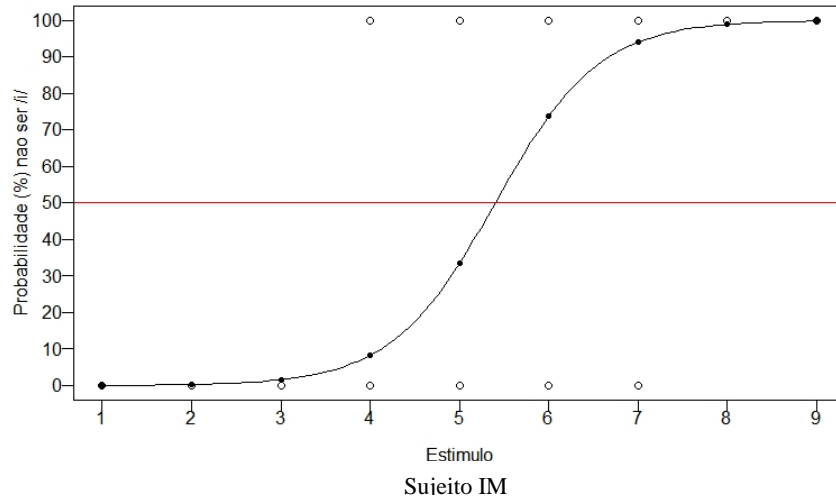


Sujeito ML

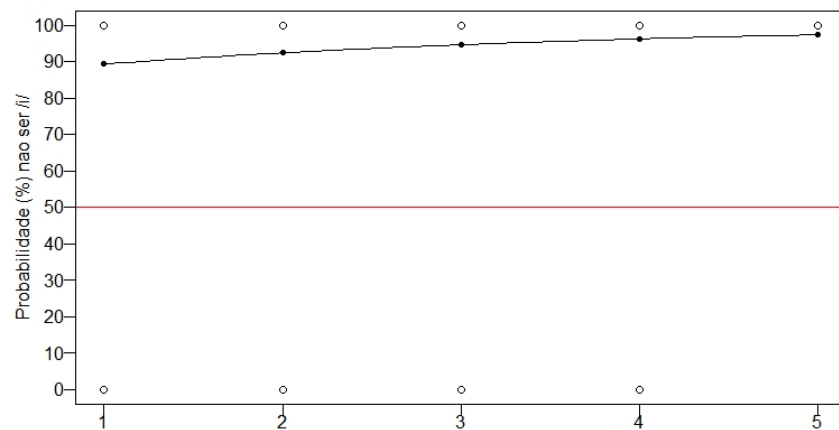


Sujeito CP

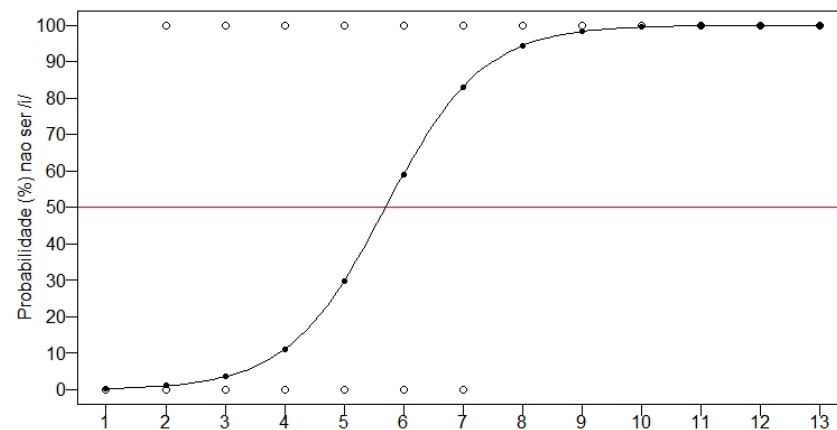
A curva de regressão logística para a transição da categoria /i/ para não /i/ (3/4)



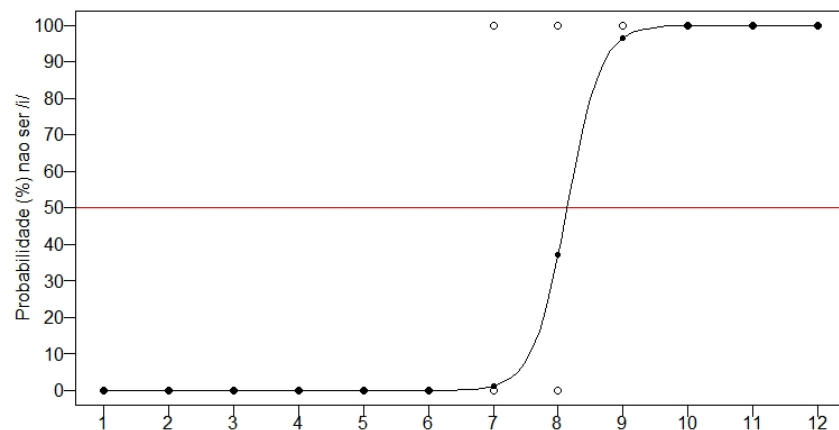
A curva de regressão logística para a transição da categoria /i/ para não /i/ (4/4)



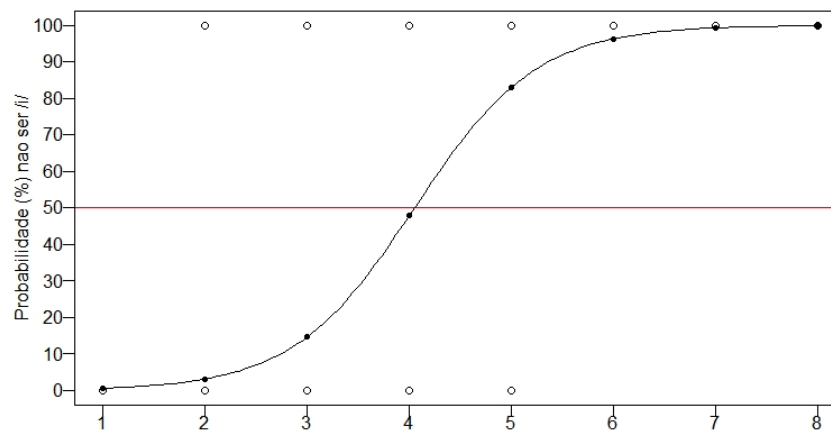
Estimulo
Sujeito AC



Estimulo
Sujeito BR

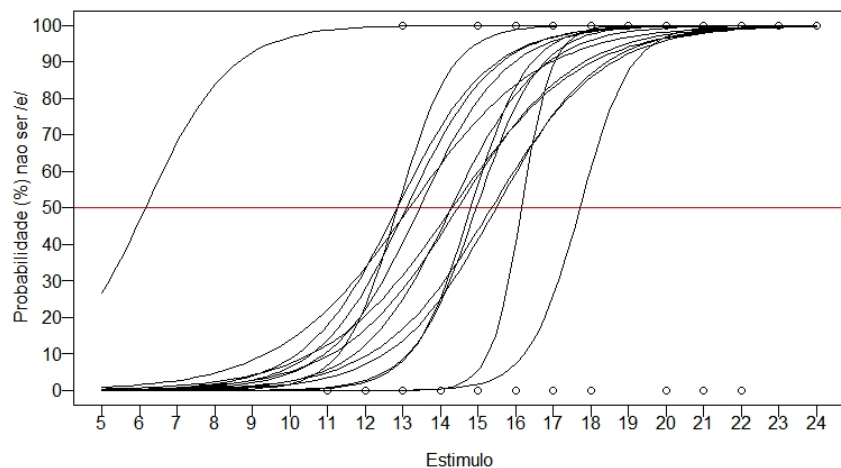


Estimulo
Sujeito CL

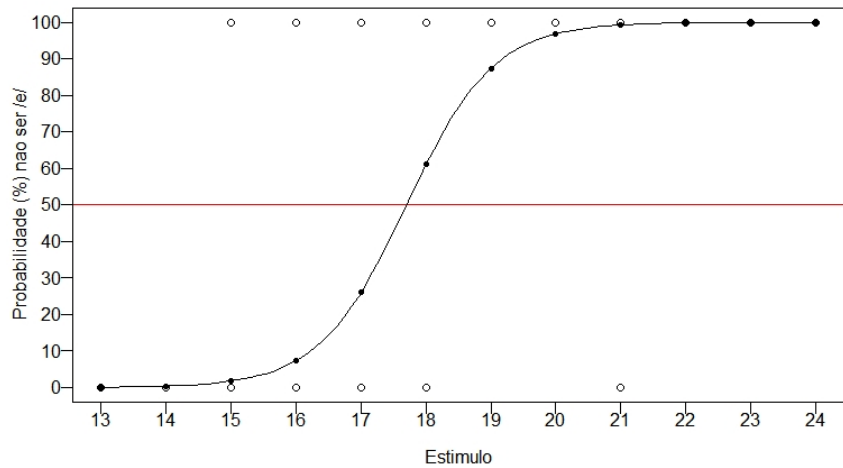


Estimulo
Sujeito LP

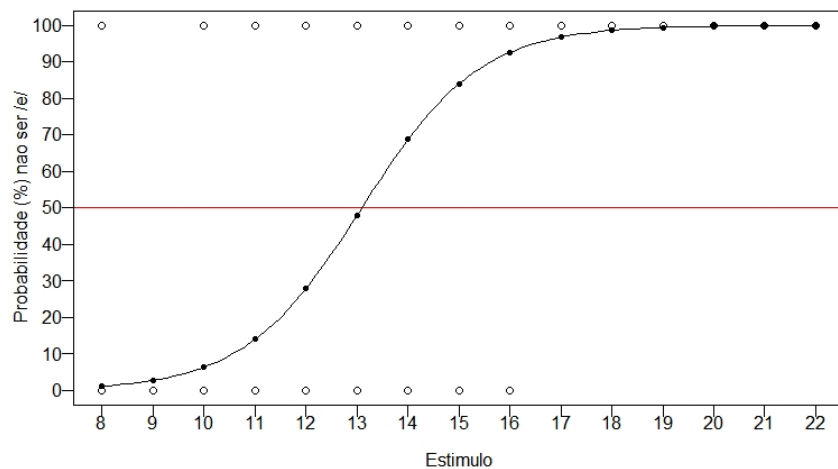
A curva de regressão logística para a transição da categoria /e/ para não /e/ (1/4)



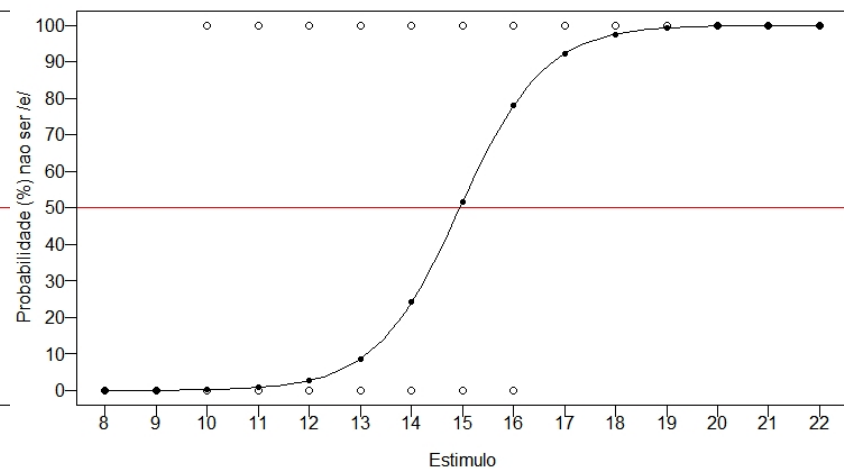
Todos os participantes



Sujeito CC

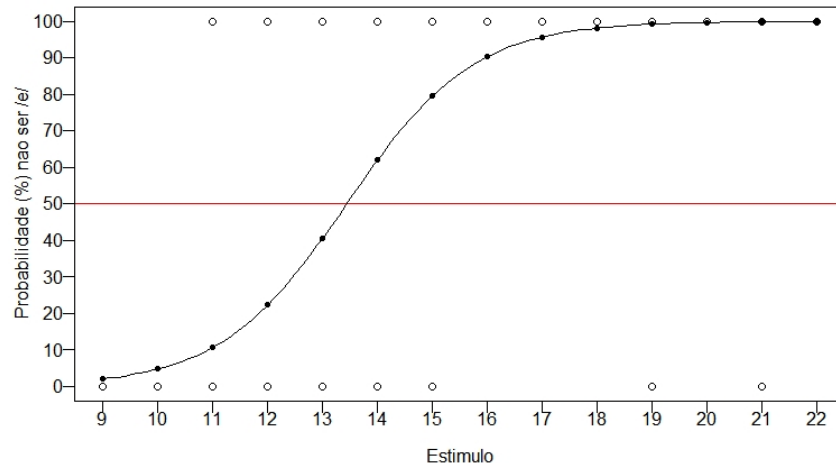


Sujeito AP

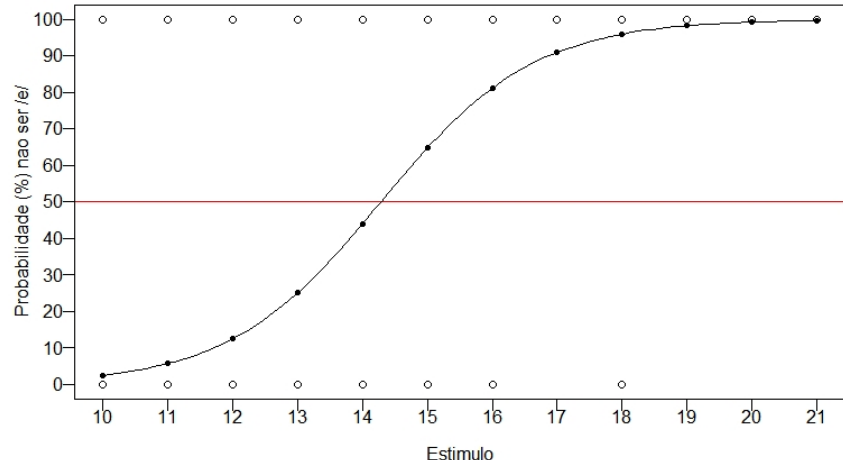


Sujeito BM

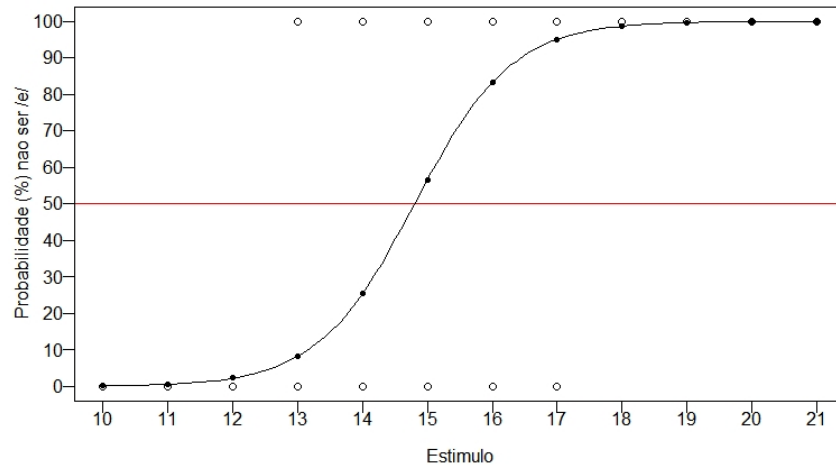
A curva de regressão logística para a transição da categoria /e/ para não /e/ (2/4)



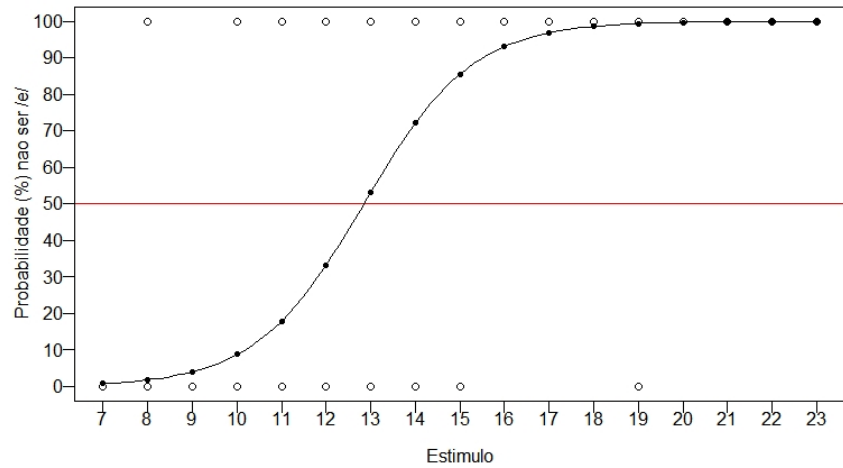
Sujeito MM



Sujeito MF

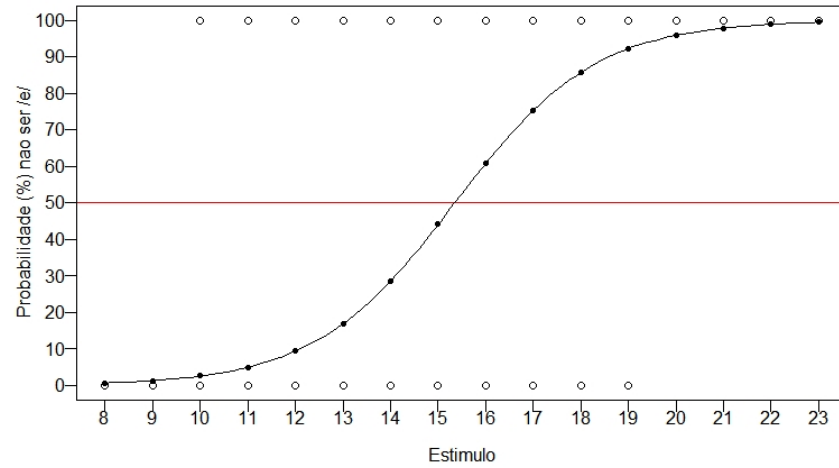


Sujeito ML

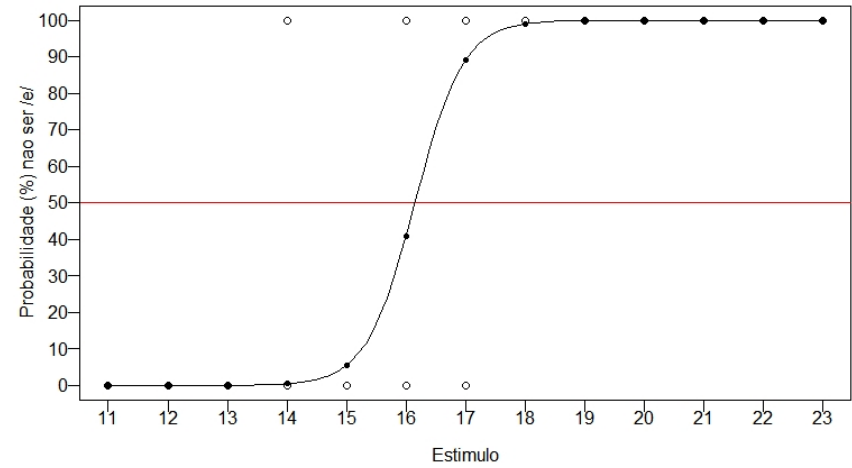


Sujeito CP

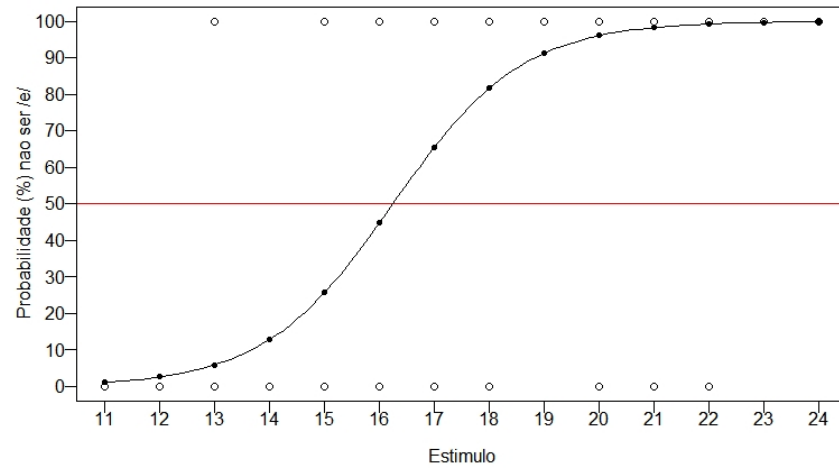
A curva de regressão logística para a transição da categoria /e/ para não /e/ (3/4)



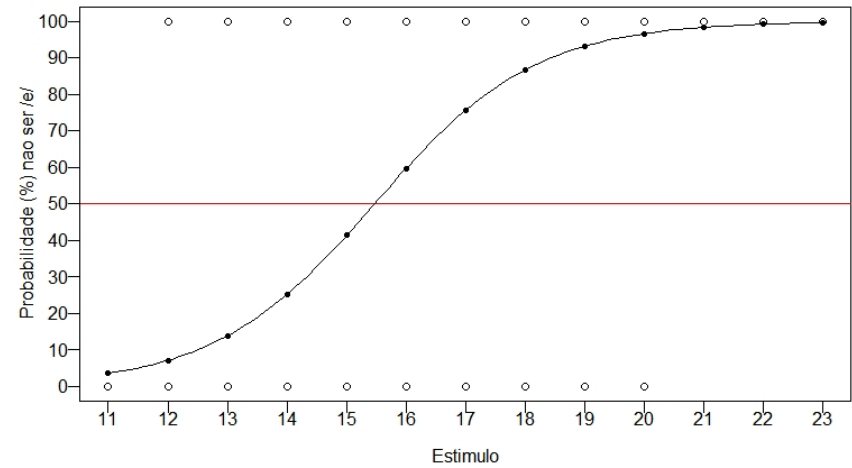
Sujeito IM



Sujeito MC

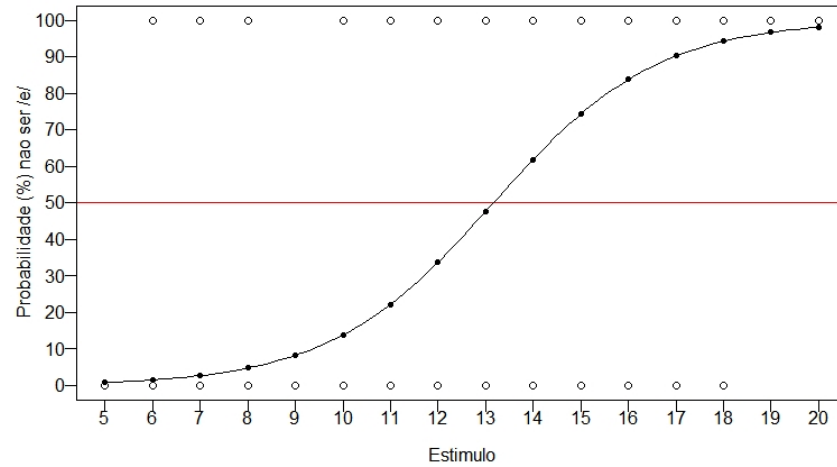


Sujeito GS

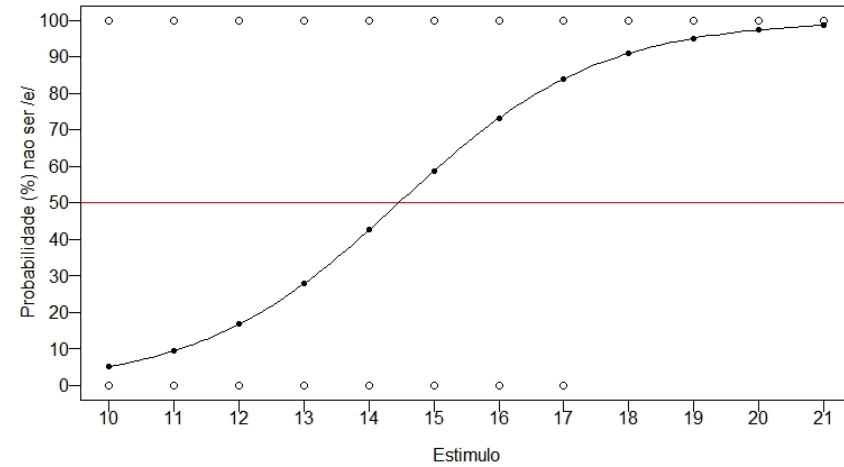


Sujeito FM

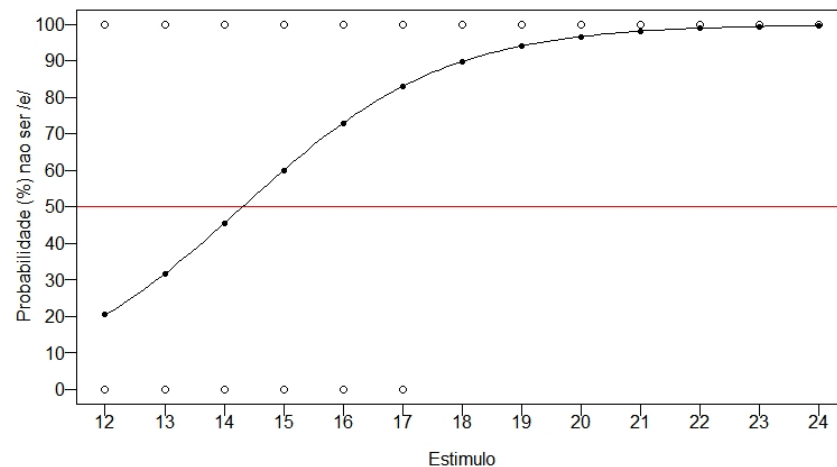
A curva de regressão logística para a transição da categoria /e/ para não /e/ (4/4)



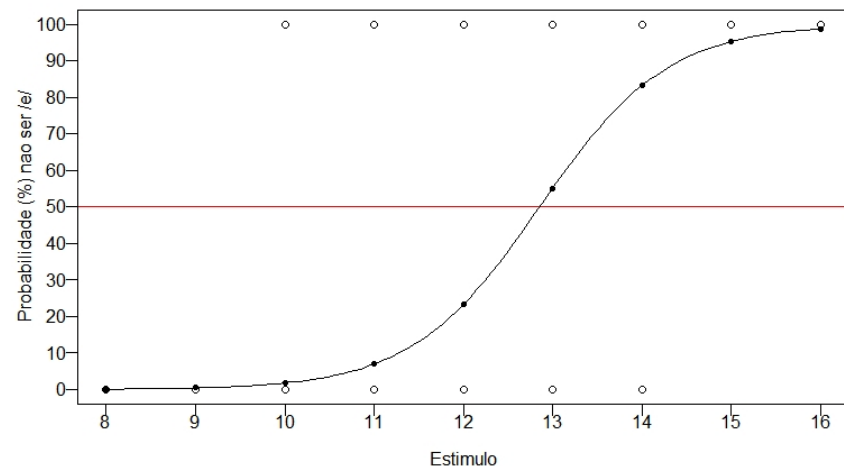
Sujeito AC



Sujeito BR

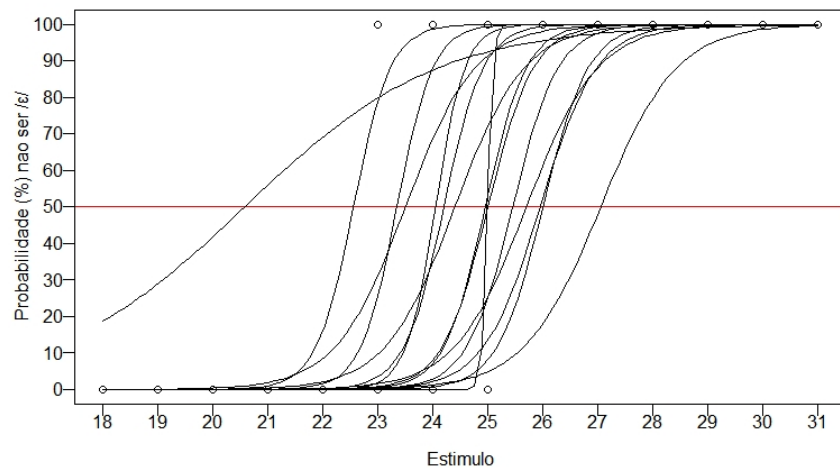


Sujeito CL

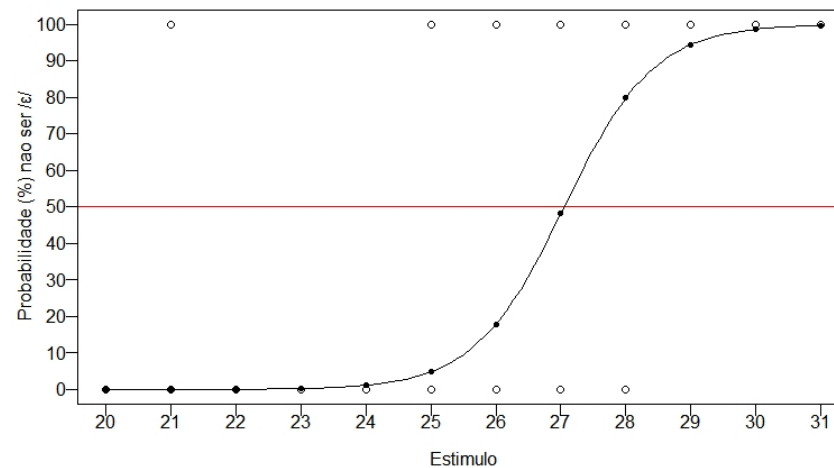


Sujeito LP

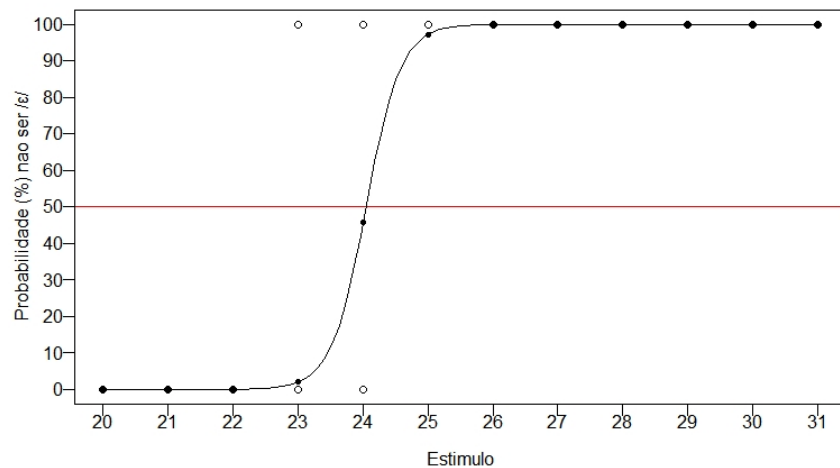
A curva de regressão logística para a transição da categoria /ε/ para não /ε/ (1/4)



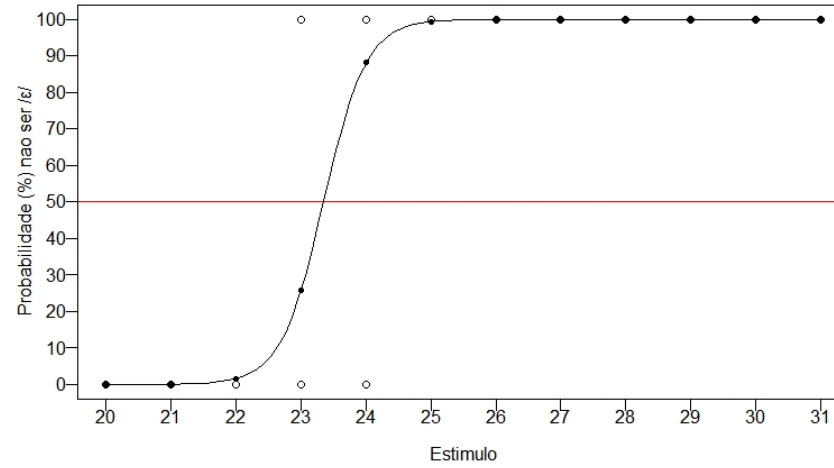
Todos os participantes



Sujeito CC

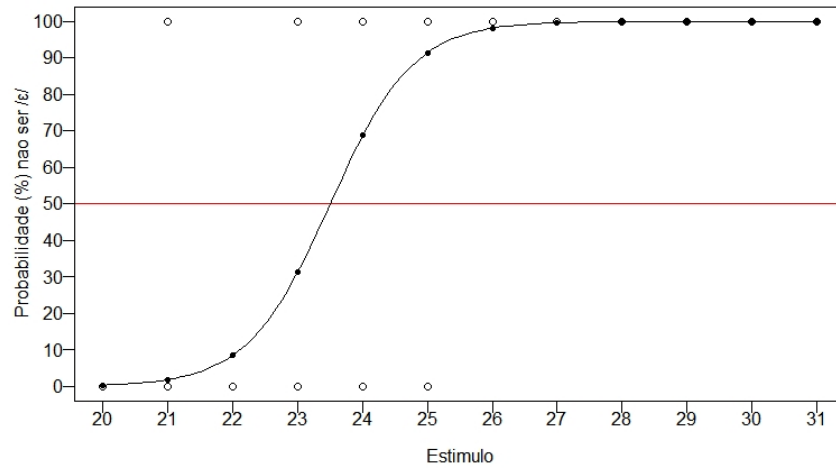


Sujeito AP

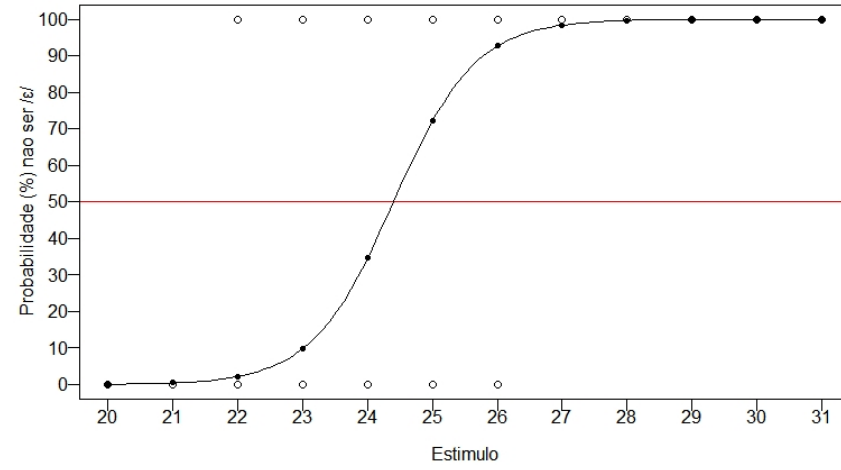


Sujeito BM

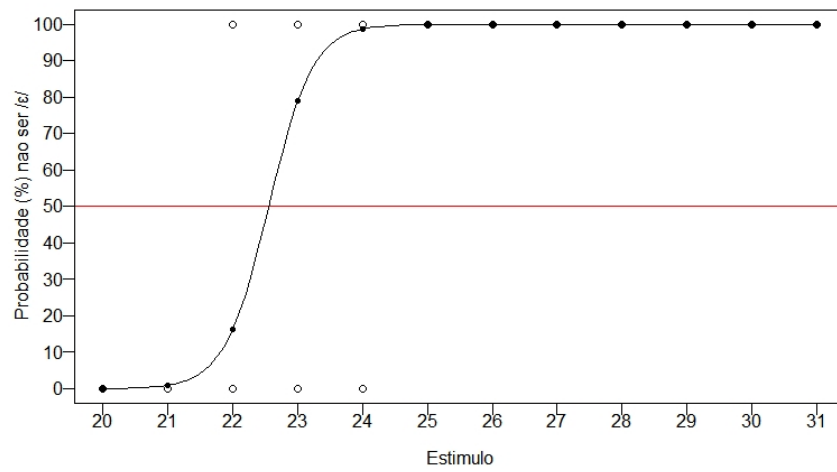
A curva de regressão logística para a transição da categoria /ε/ para não /ε/ (2/4)



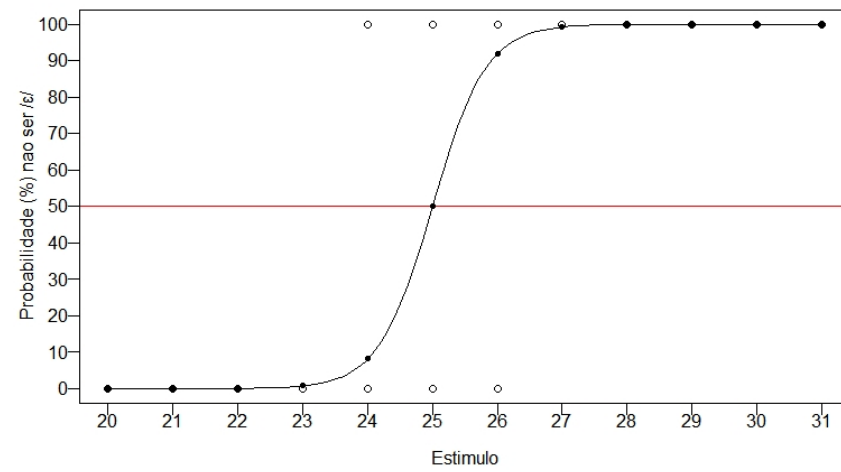
Sujeito MM



Sujeito MF

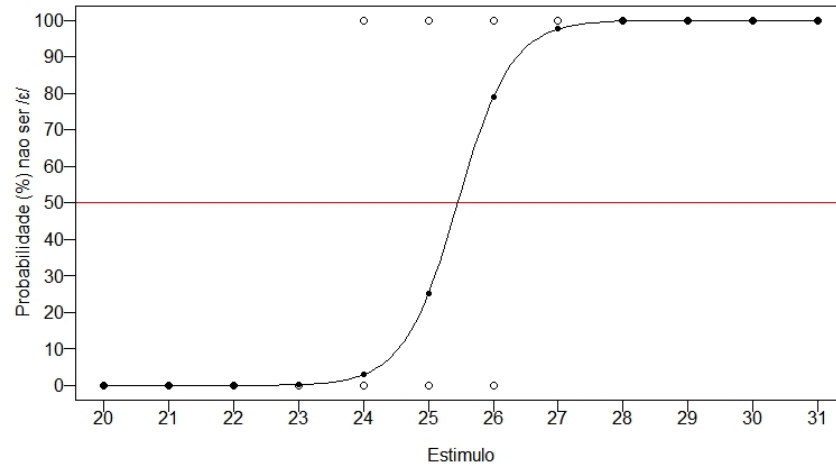


Sujeito ML

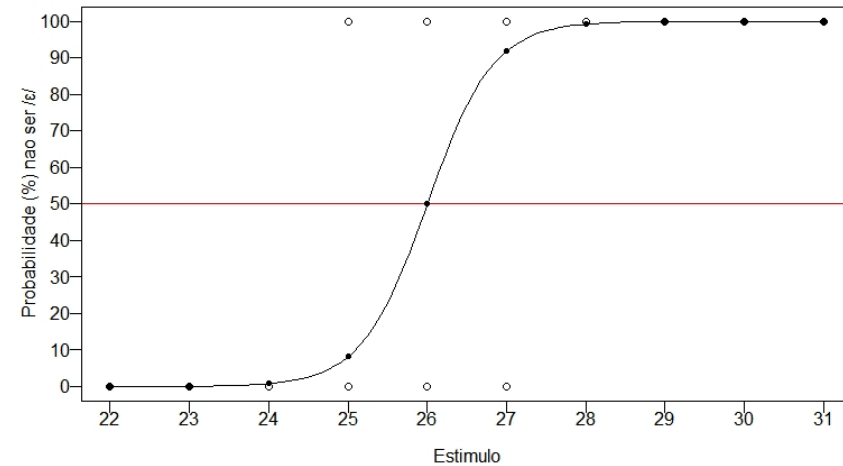


Sujeito CP

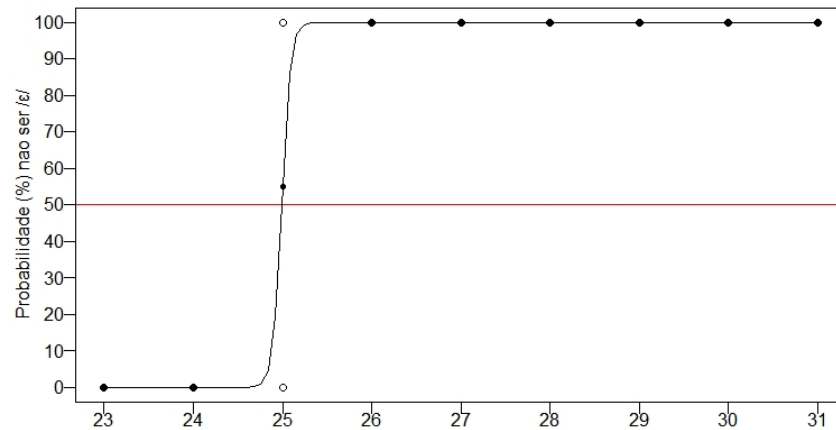
A curva de regressão logística para a transição da categoria /ε/ para não /ε/ (3/4)



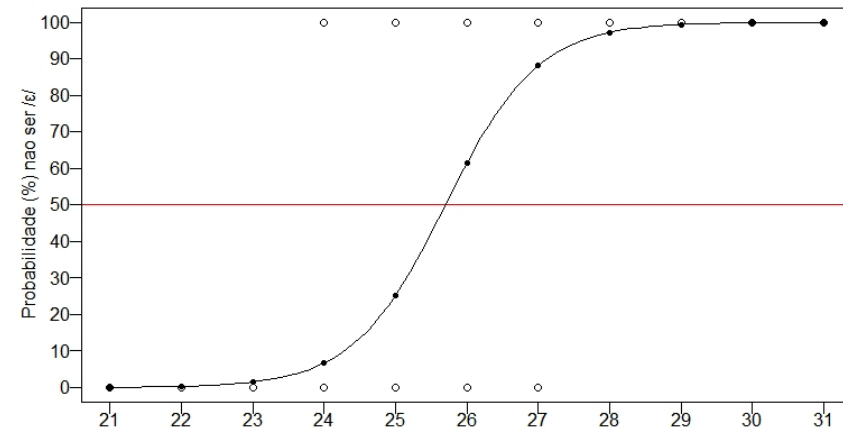
Sujeito IM



Sujeito MC

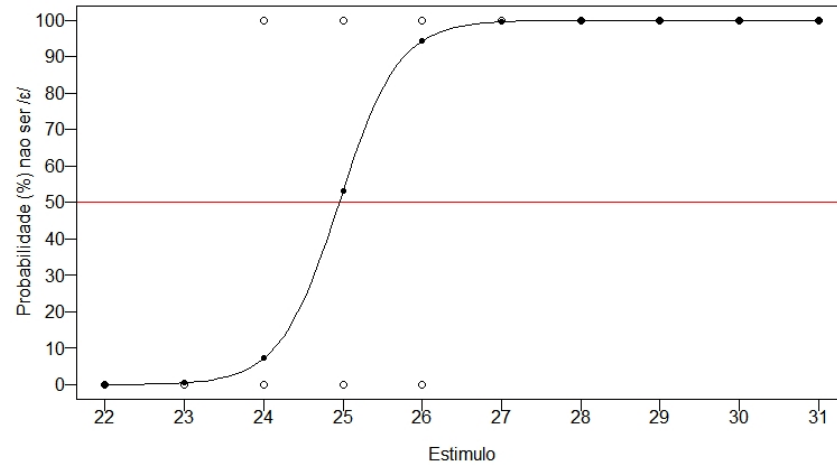


Sujeito GS

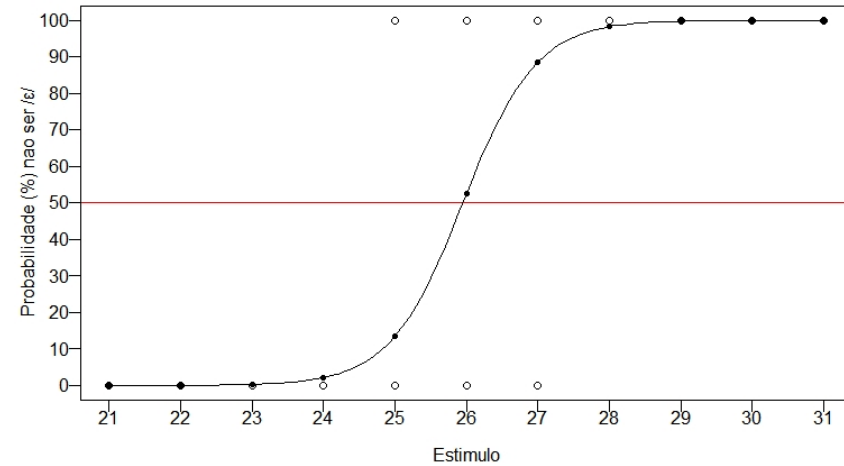


Sujeito FM

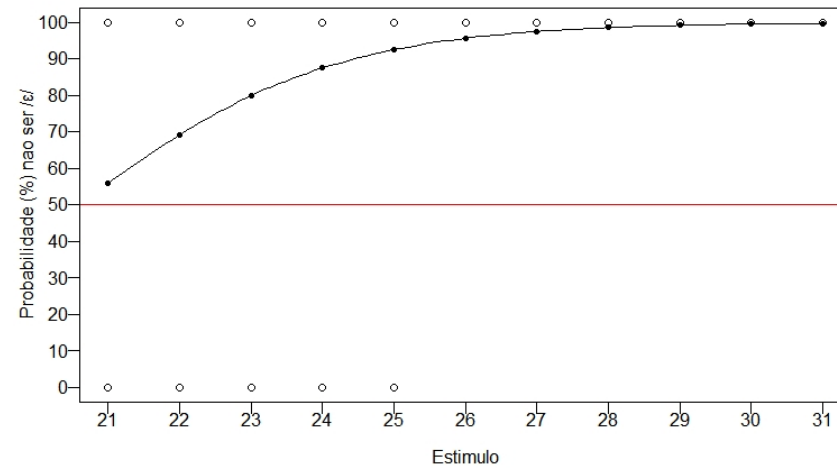
A curva de regressão logística para a transição da categoria /ε/ para não /ε/ (4/4)



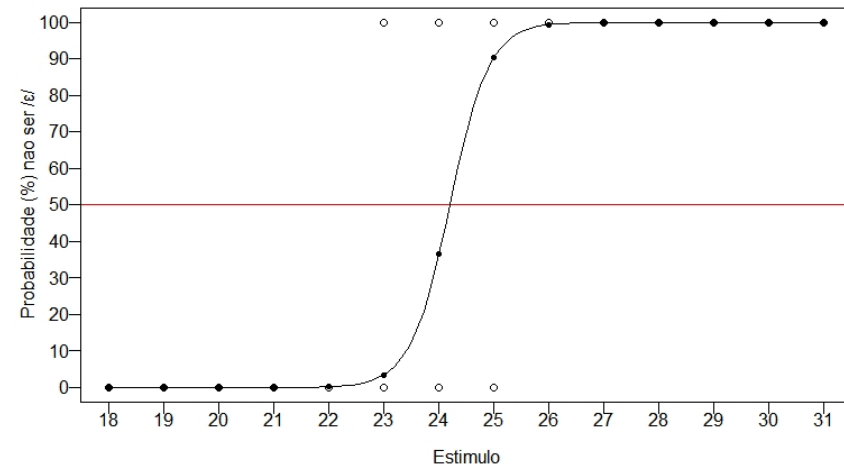
Sujeito AC



Sujeito BR

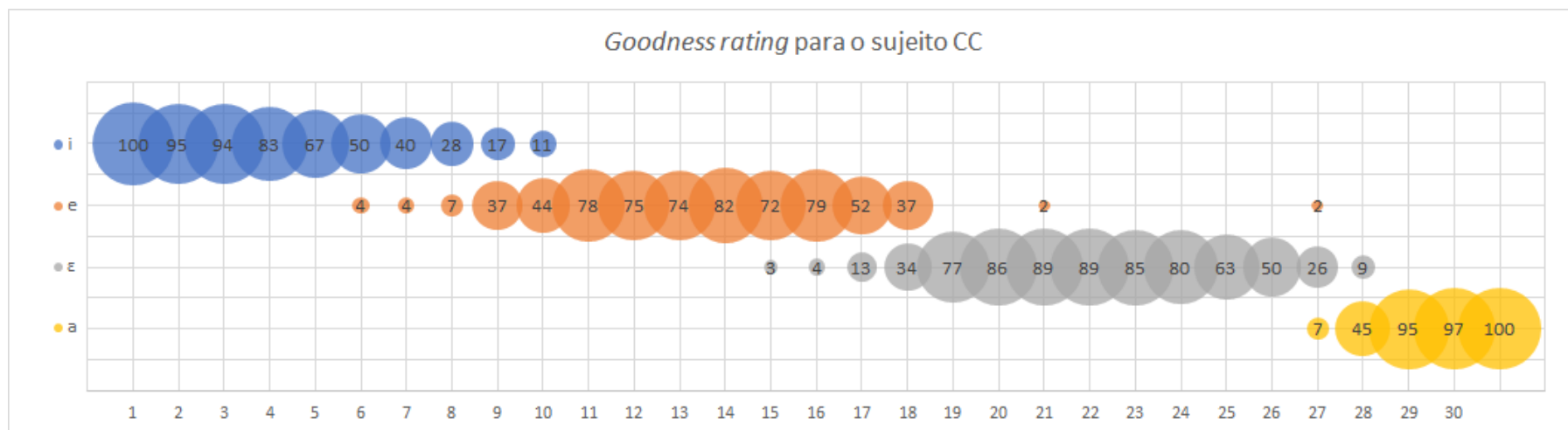
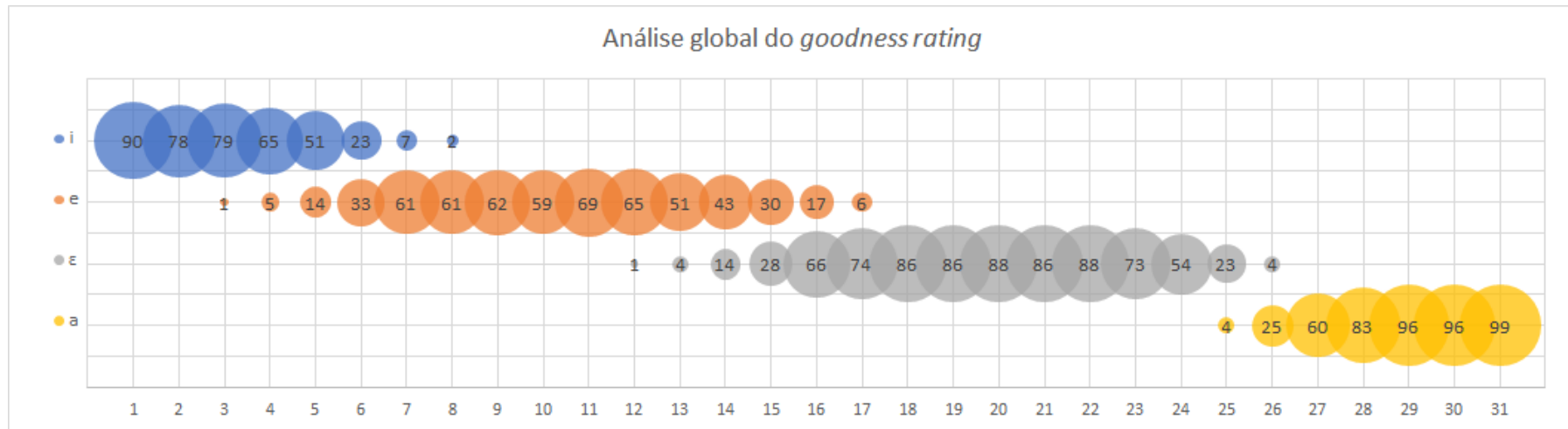


Sujeito CL

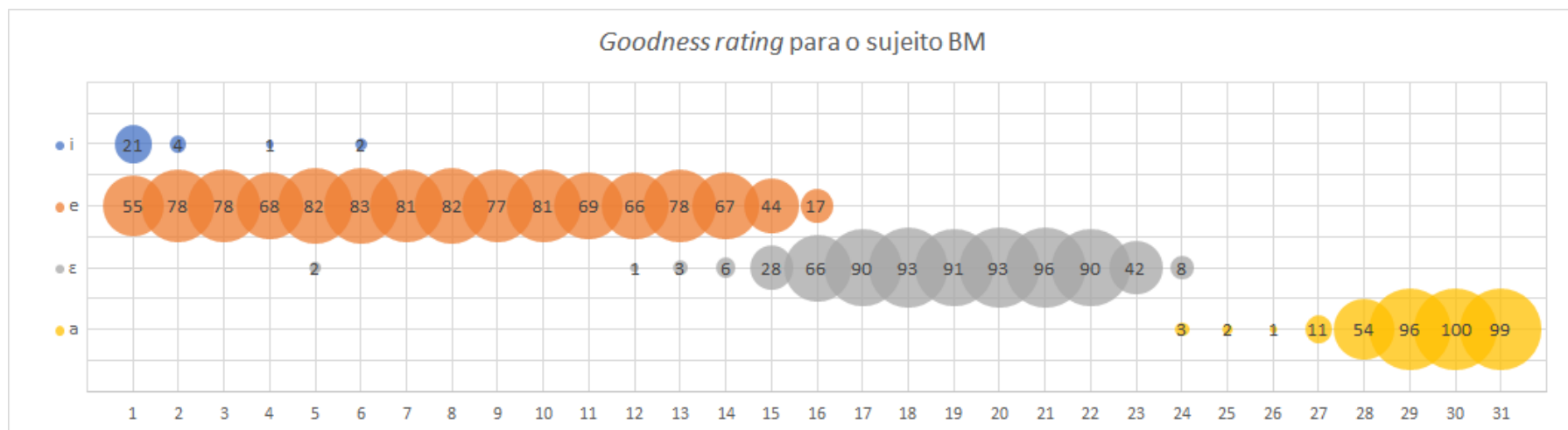
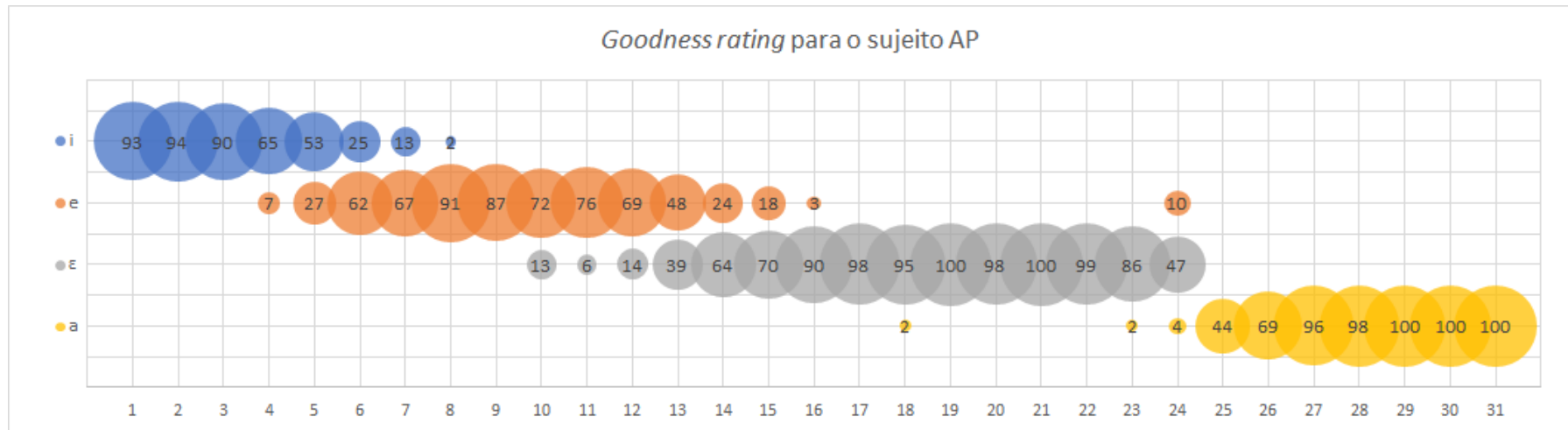


Sujeito LP

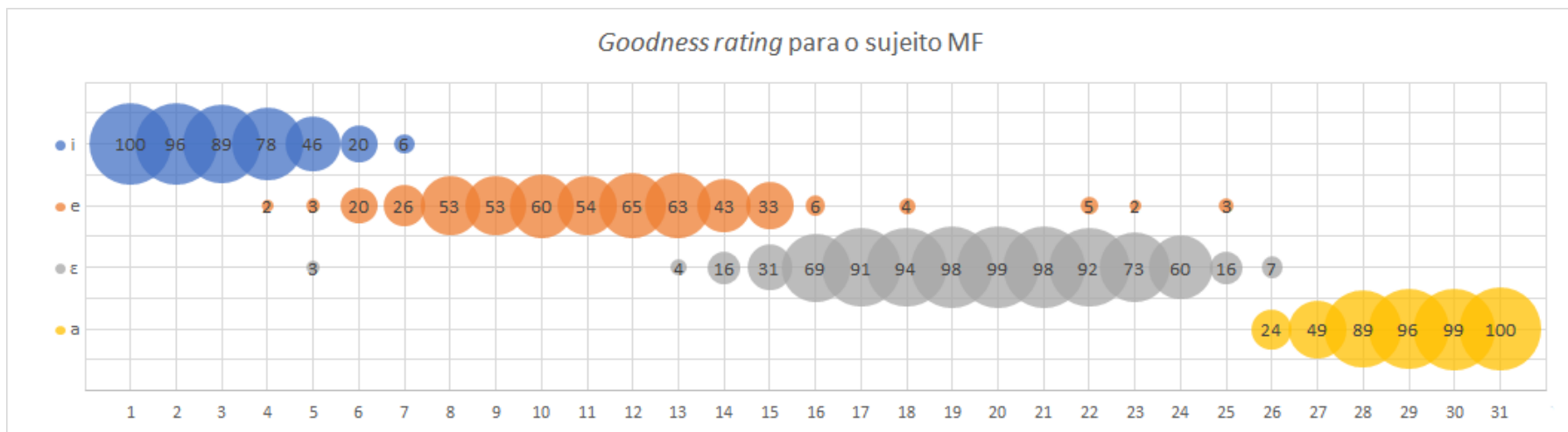
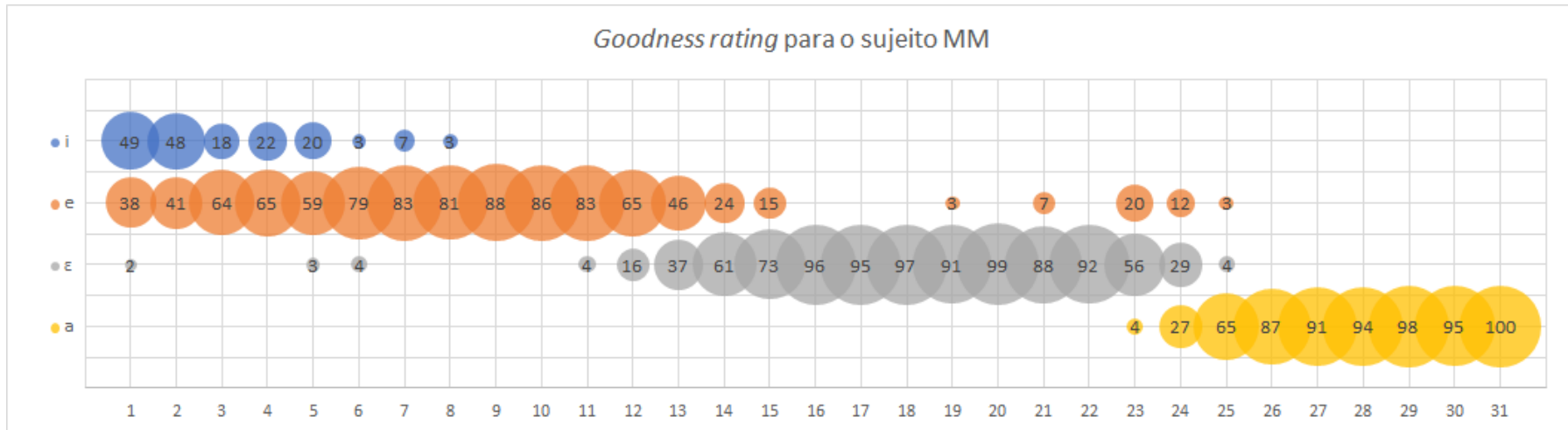
Goodness rating (1/8)



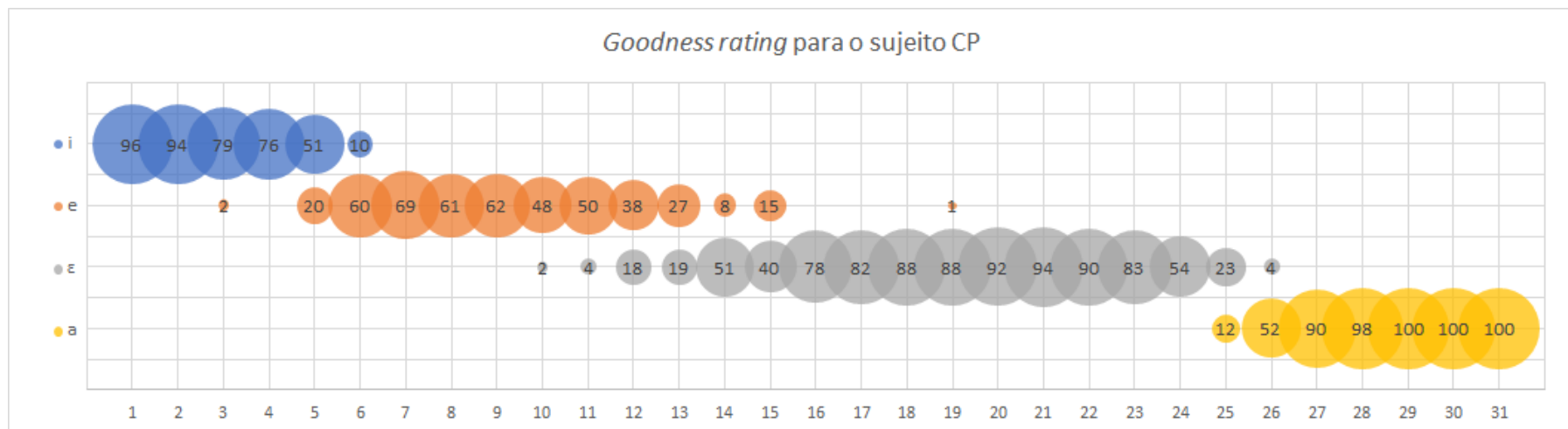
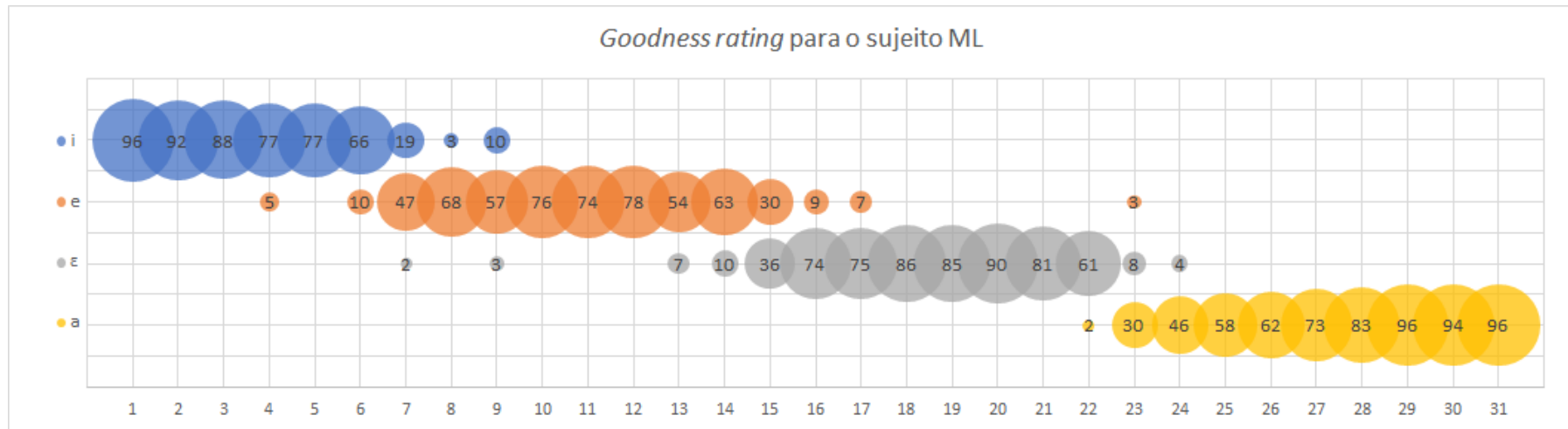
Goodness rating (2/8)



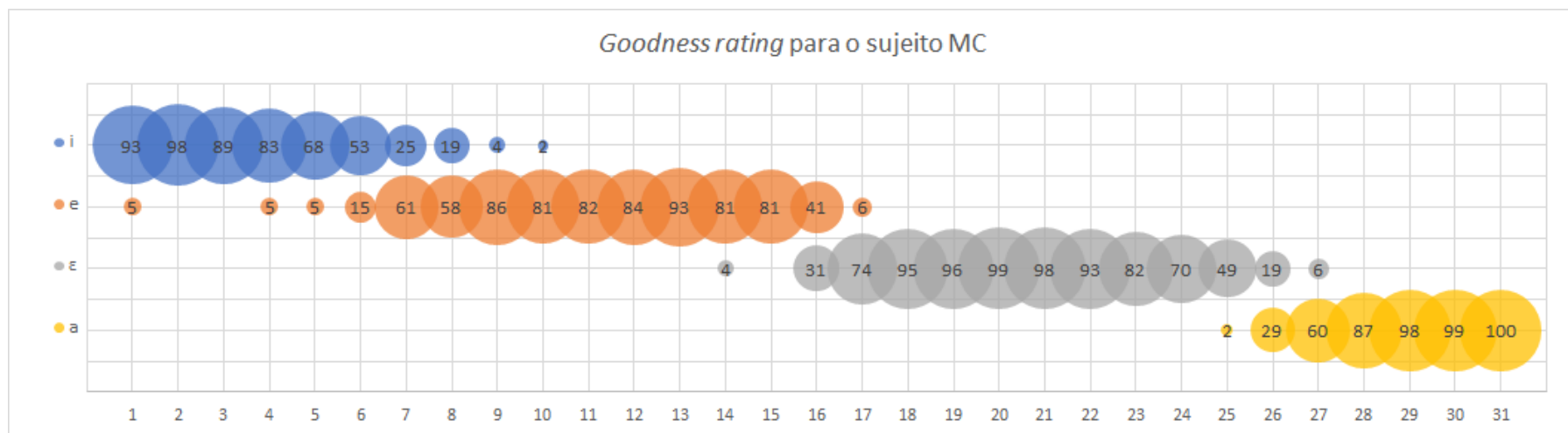
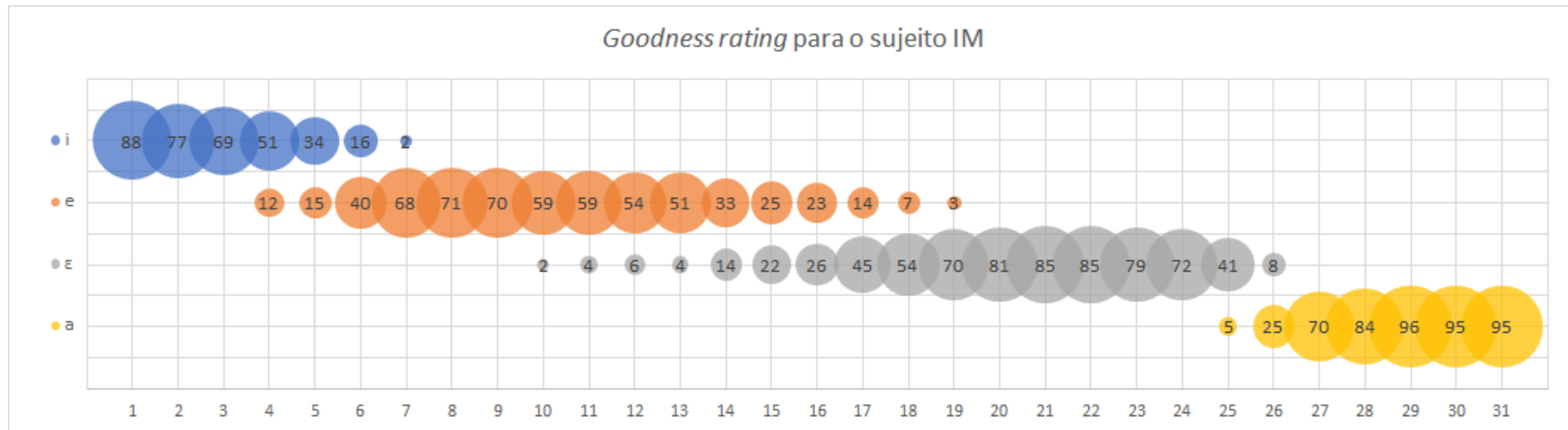
Goodness rating (3/8)



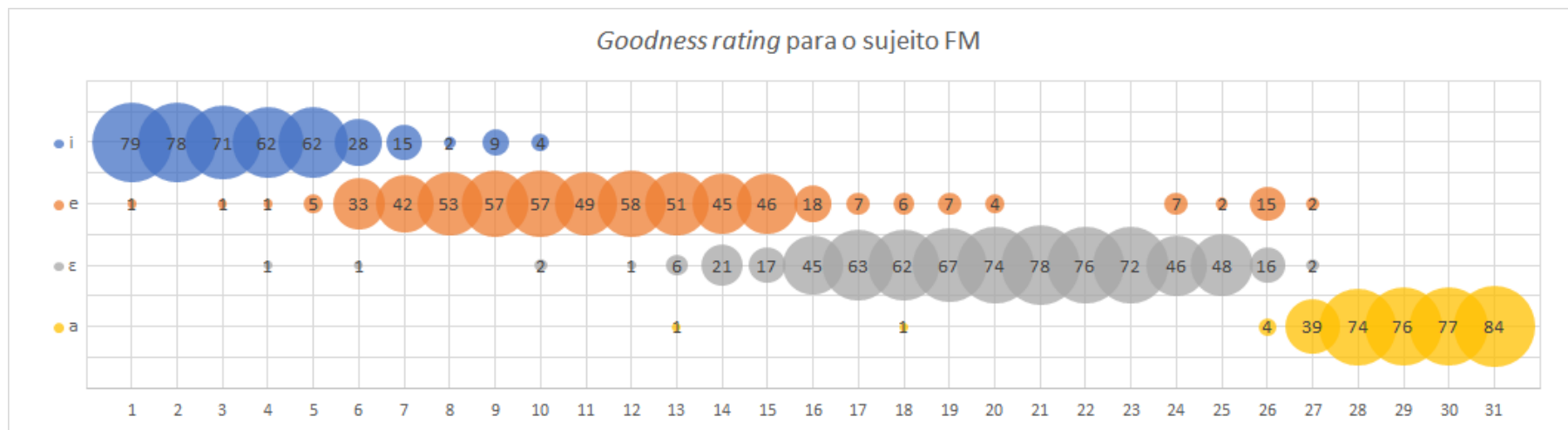
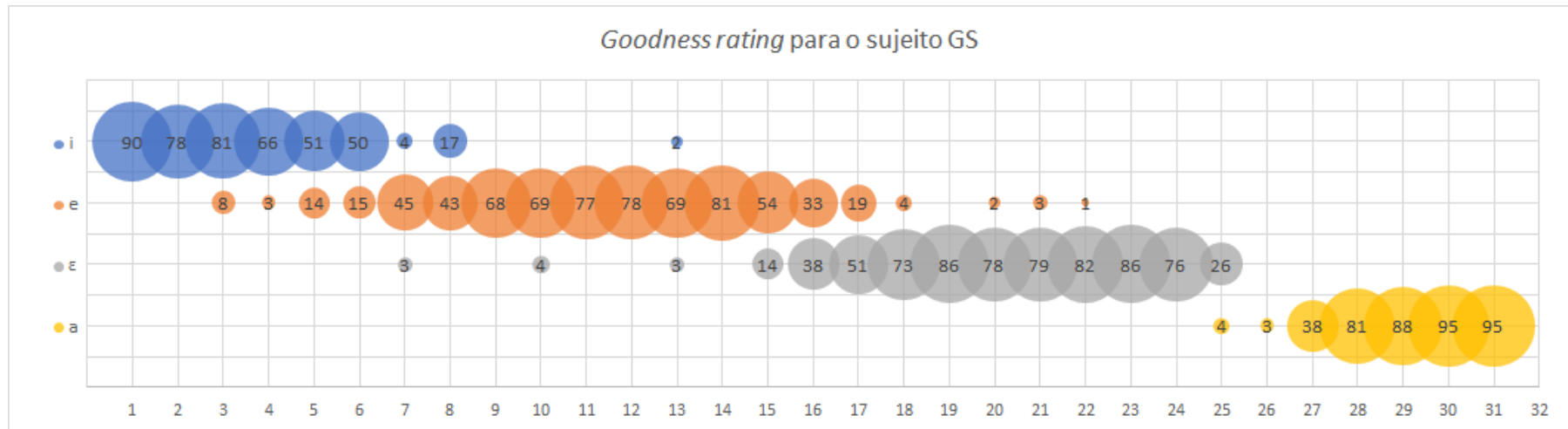
Goodness rating (4/8)



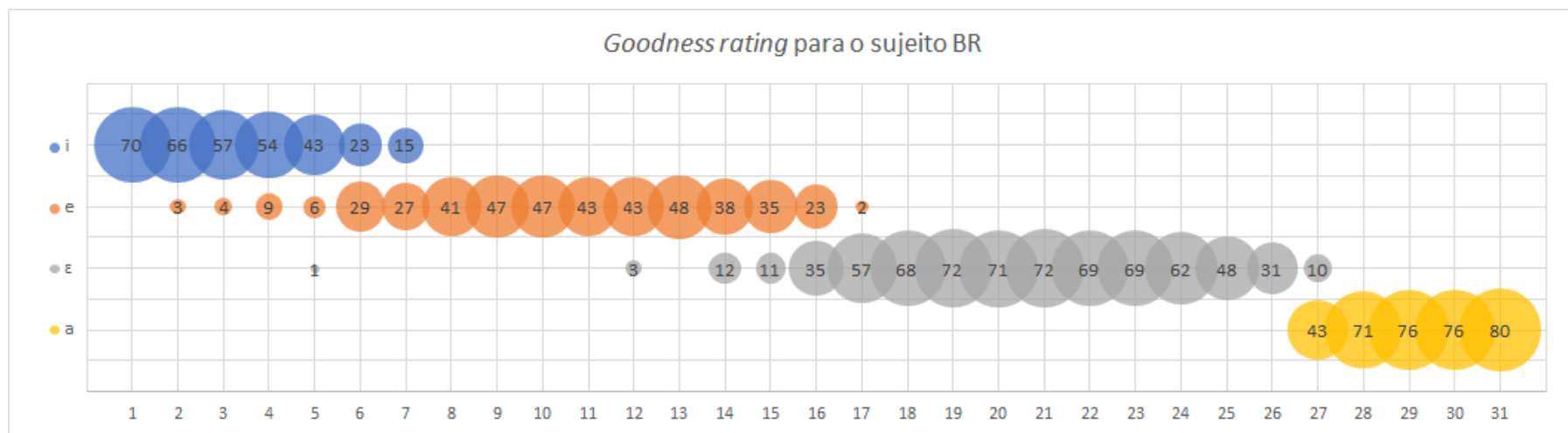
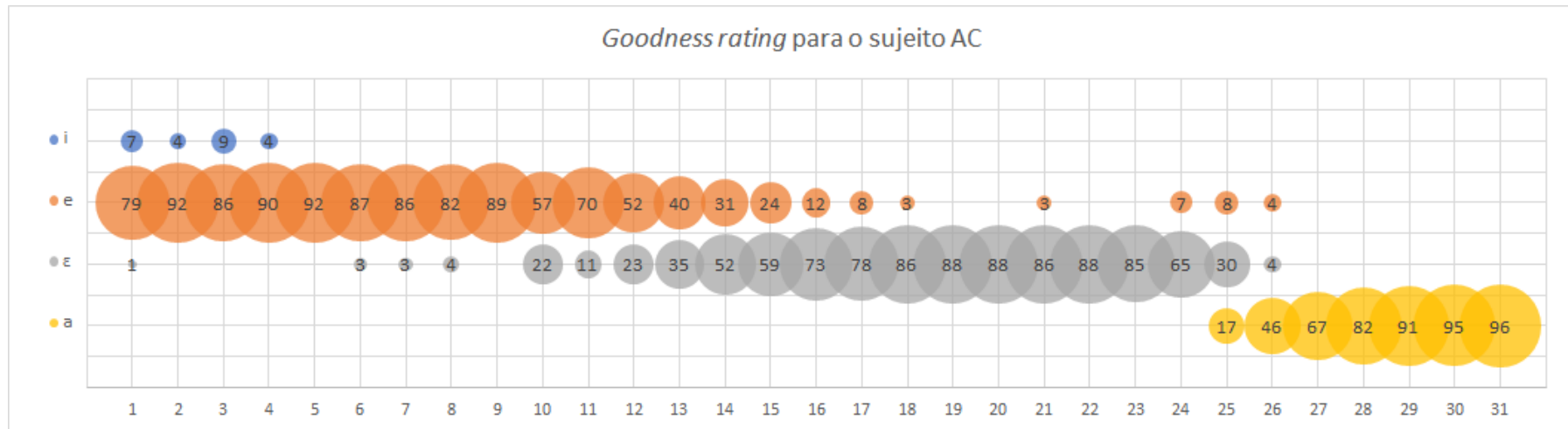
Goodness rating (5/8)



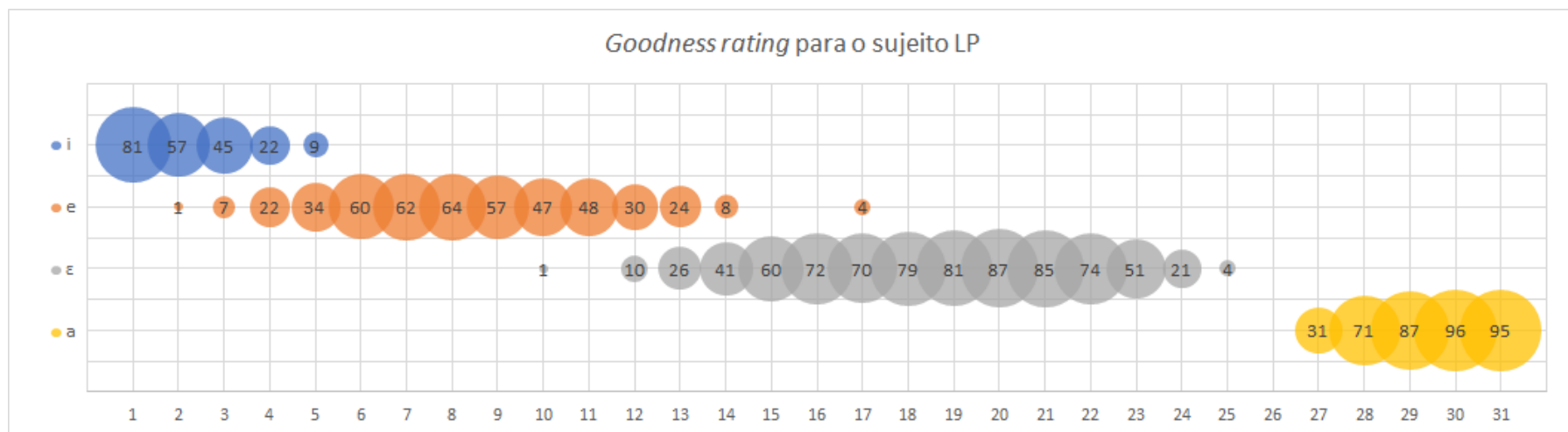
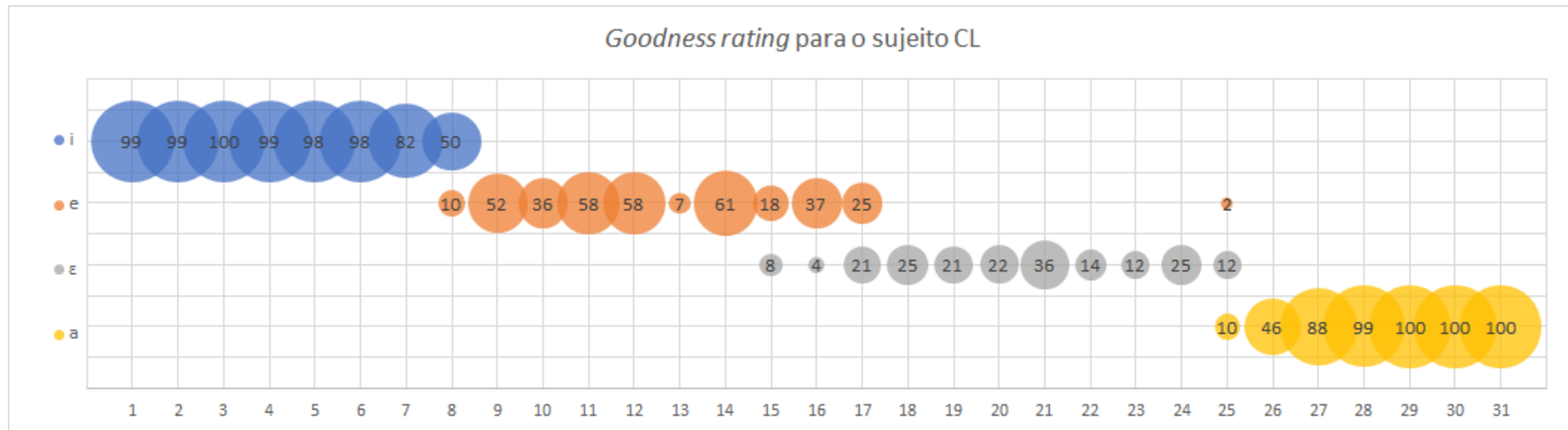
Goodness rating (6/8)



Goodness rating (7/8)

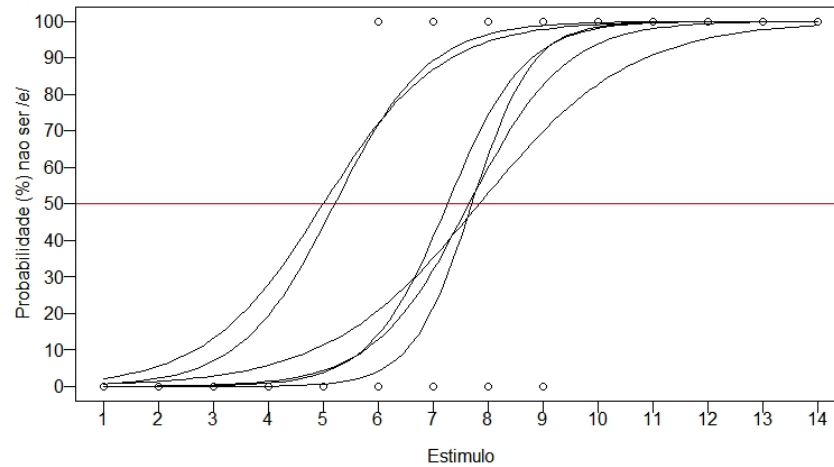


Goodness rating (8/8)

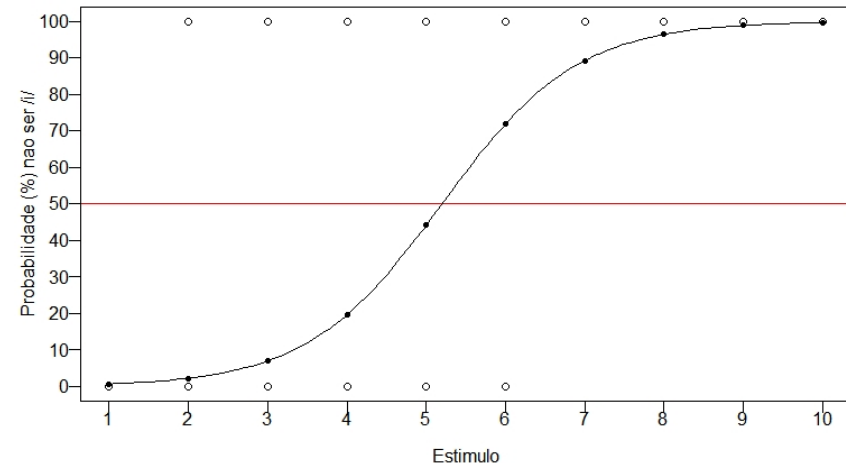


Anexo6 – Gráficos: Pré-teste

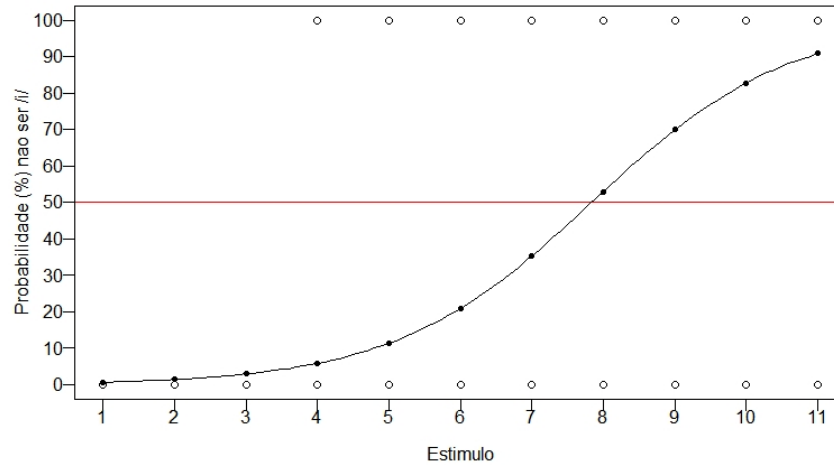
A curva de regressão logística para a transição da categoria /i/ para não /i/ no pré-teste (1/2)



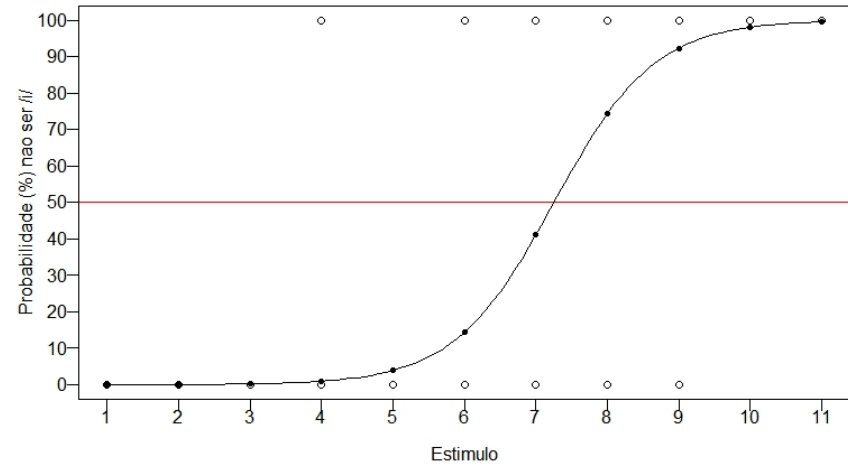
Todos os participantes do pré-teste



Sujeito AB

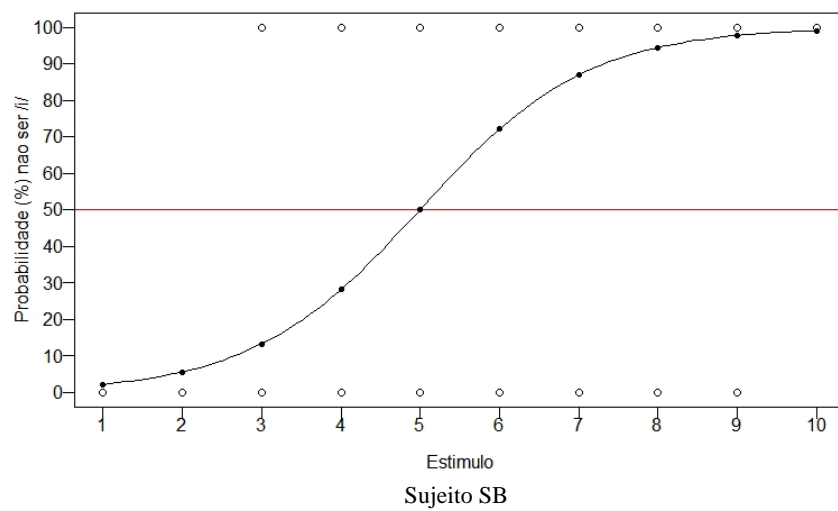
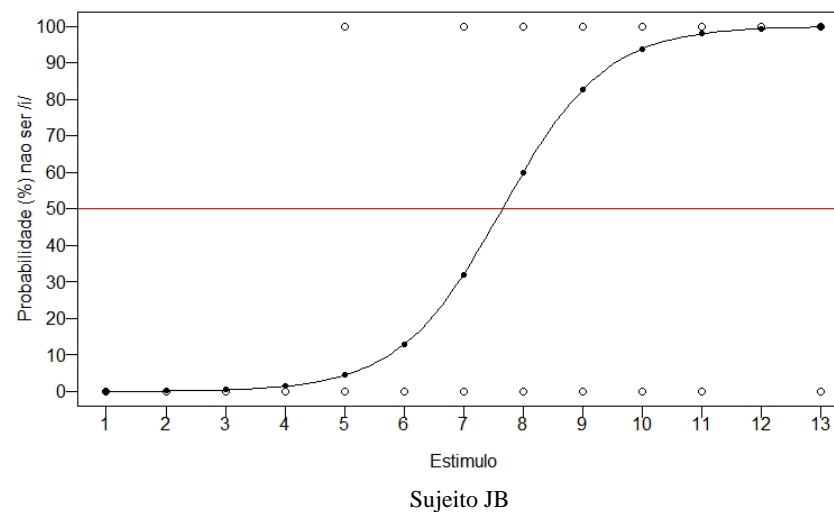
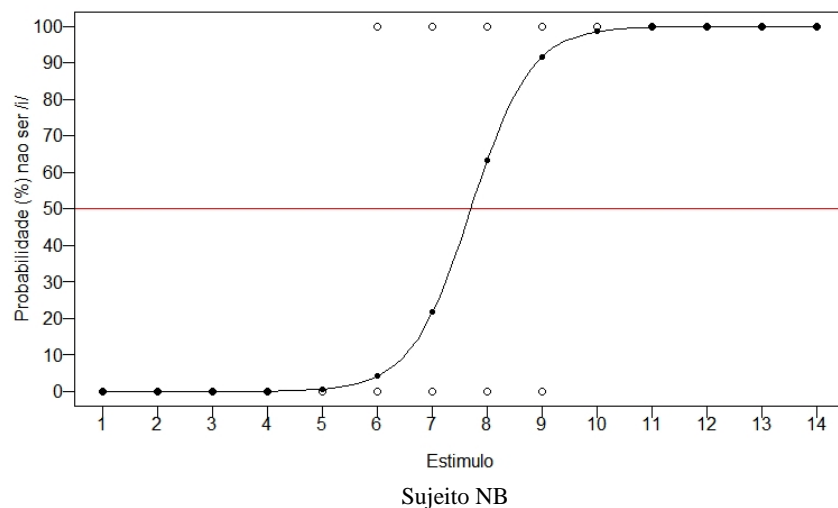


Sujeito VB

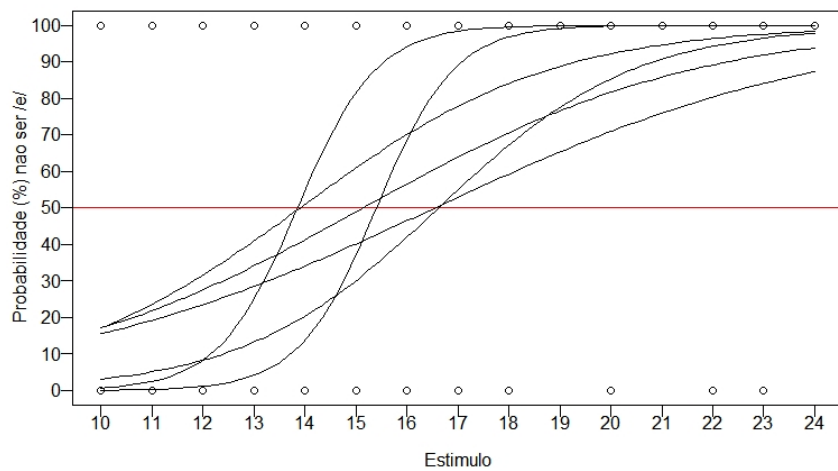


Sujeito FB

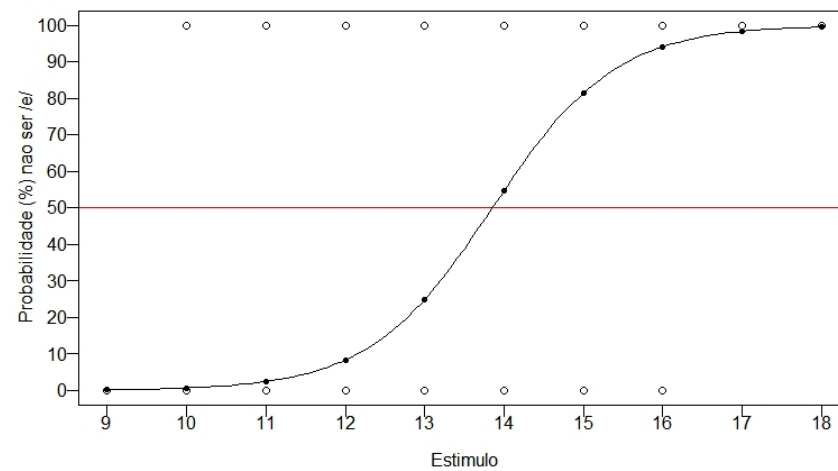
A curva de regressão logística para a transição da categoria /i/ para não /i/ no pré-teste (2/2)



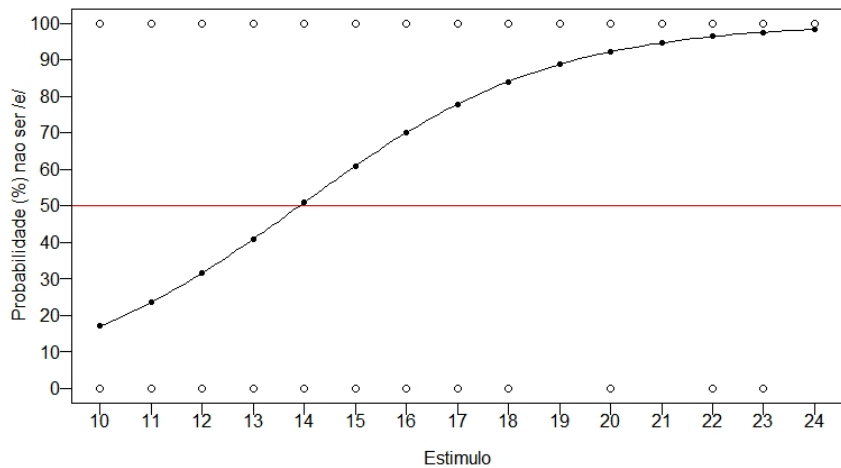
A curva de regressão logística para a transição da categoria /e/ para não /e/ no pré-teste (1/2)



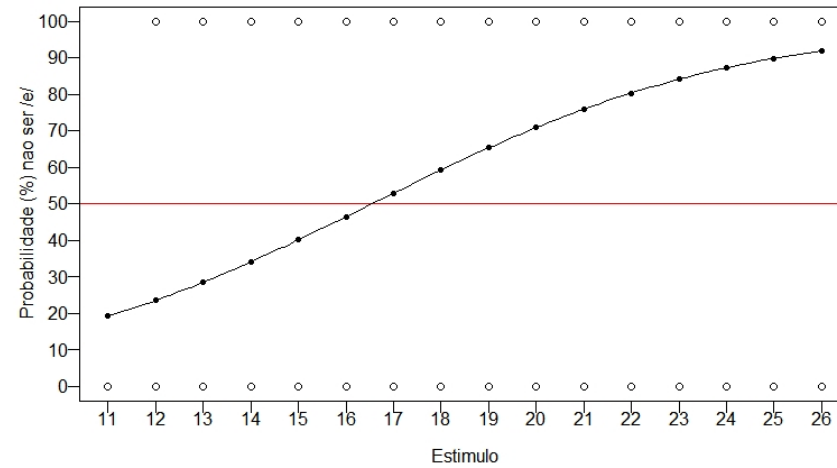
Todos os participantes do pré-teste



Sujeito AB

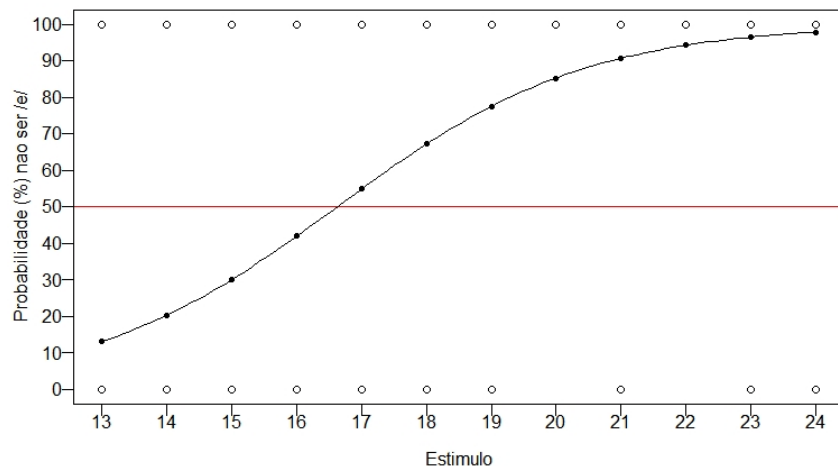


Sujeito VB

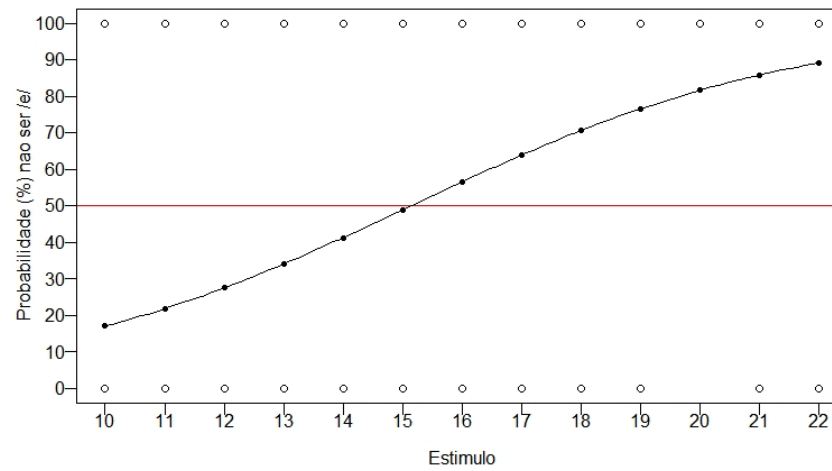


Sujeito FB

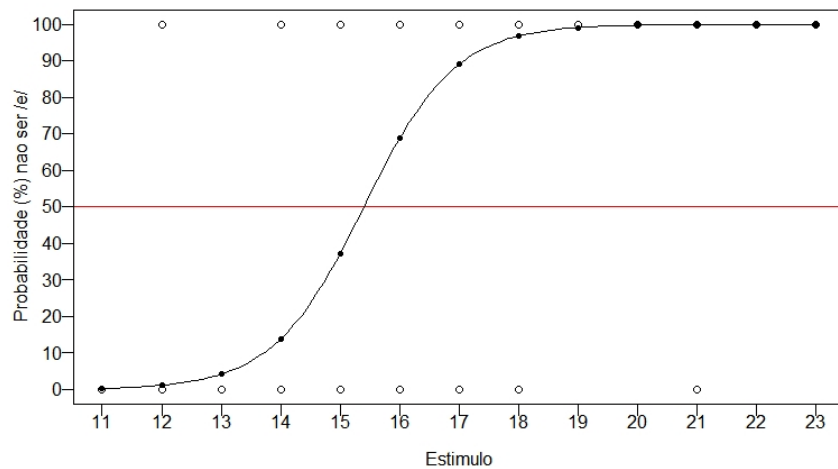
A curva de regressão logística para a transição da categoria /e/ para não /e/ no pré-teste (2/2)



Sujeito NB

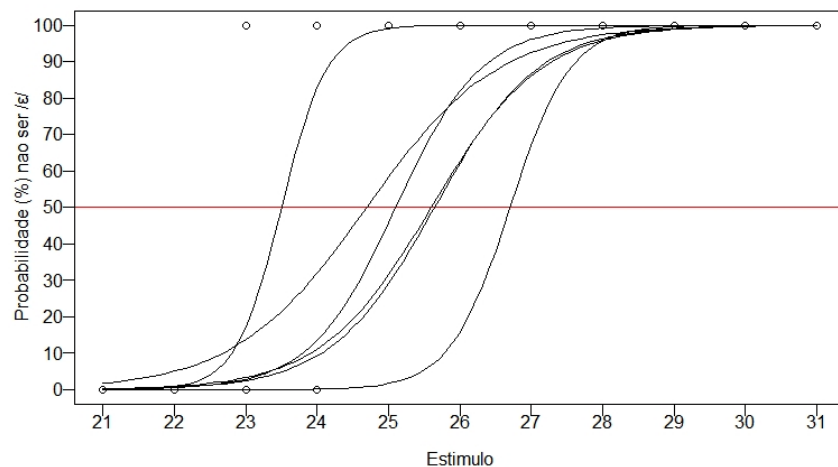


Sujeito JB

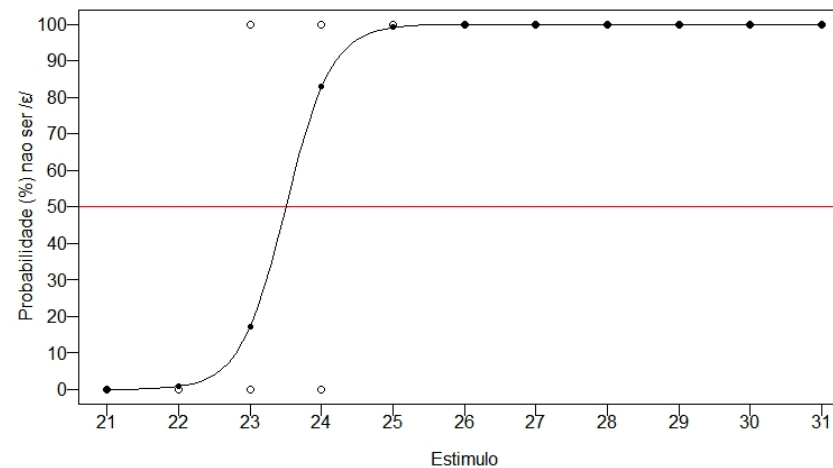


Sujeito SB

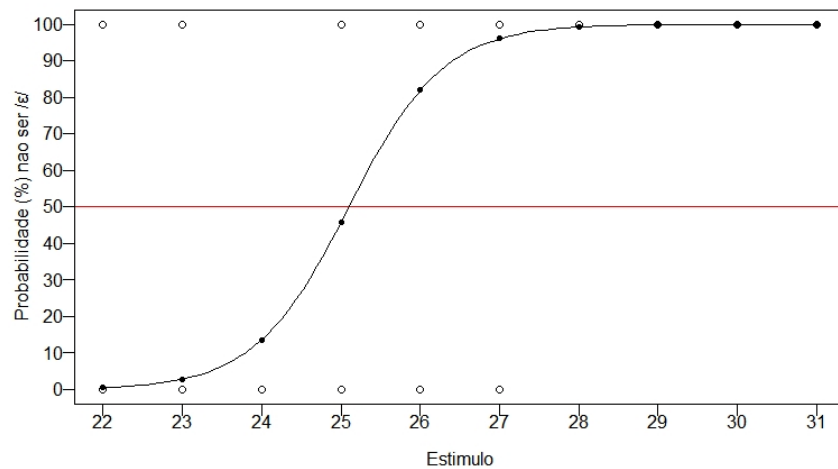
A curva de regressão logística para a transição da categoria /ε/ para não /ε/ no pré-teste (1/2)



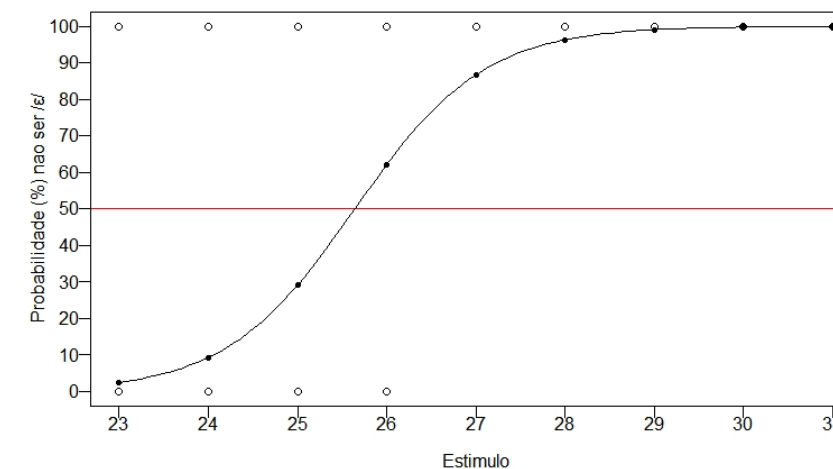
Todos os participantes do pré-teste



Sujeito AB

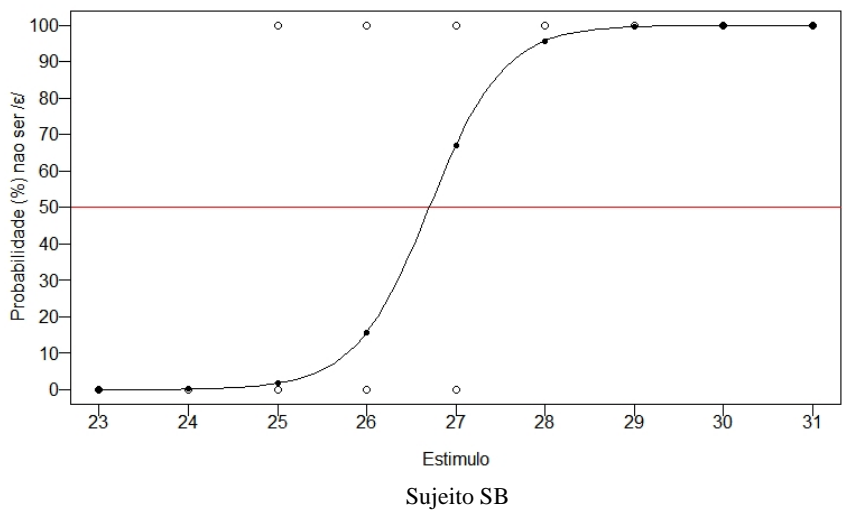
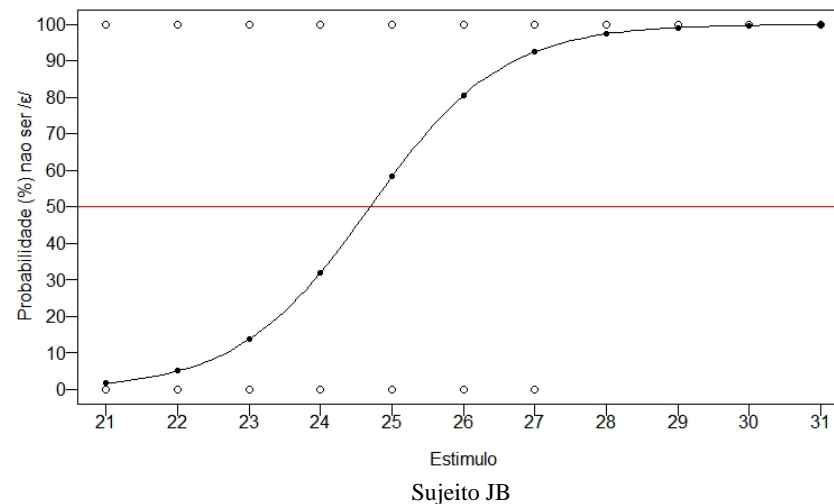
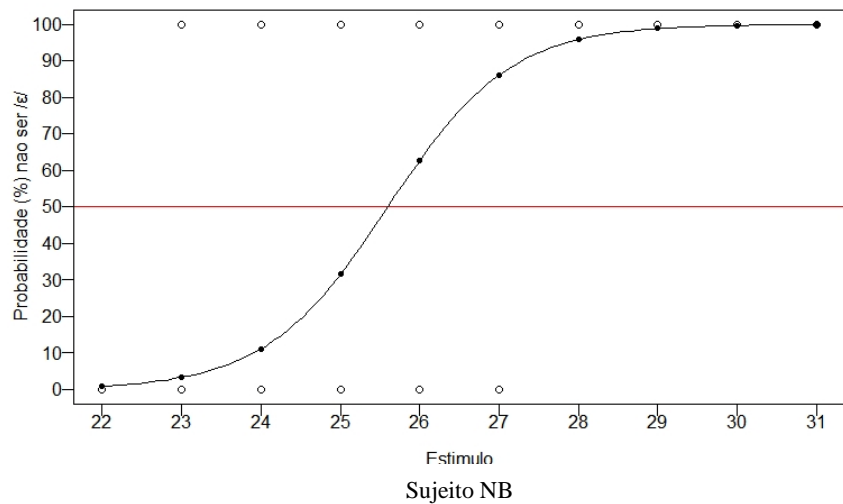


Sujeito VB

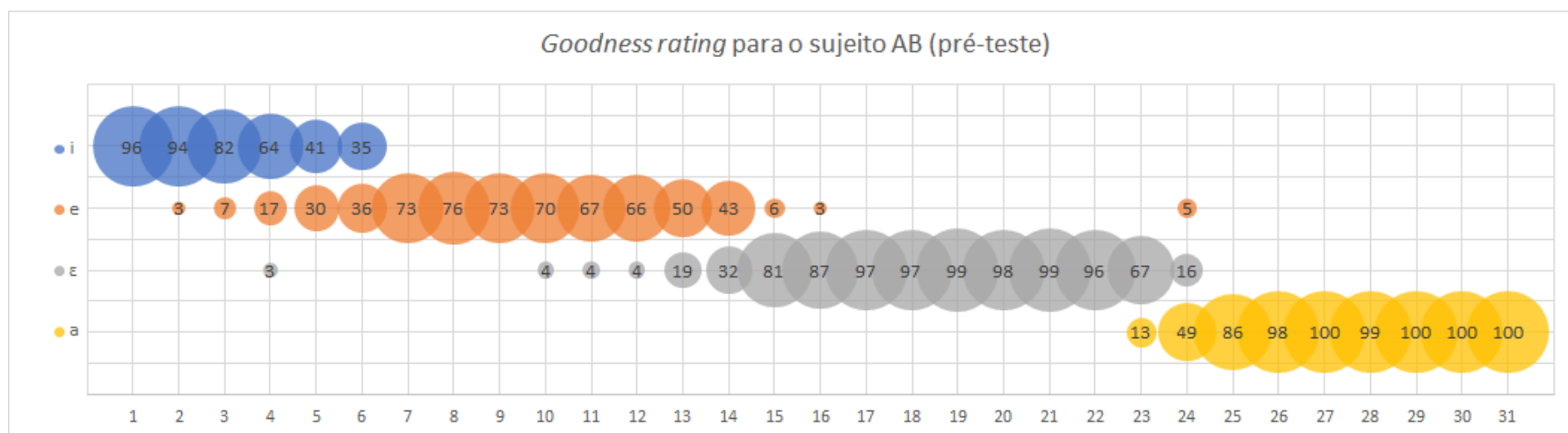
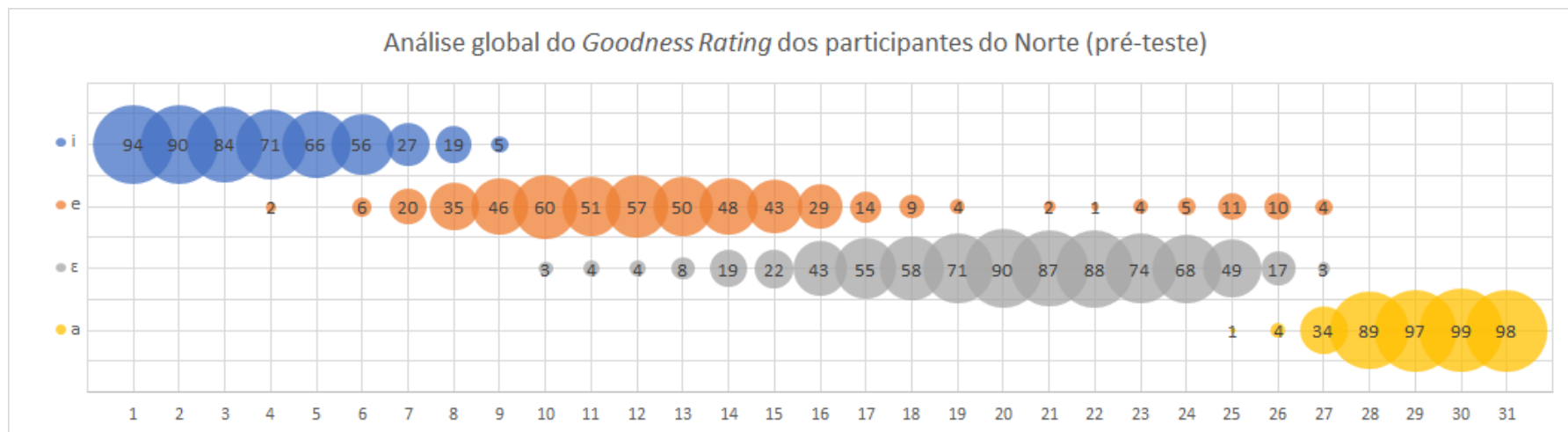


Sujeito FB

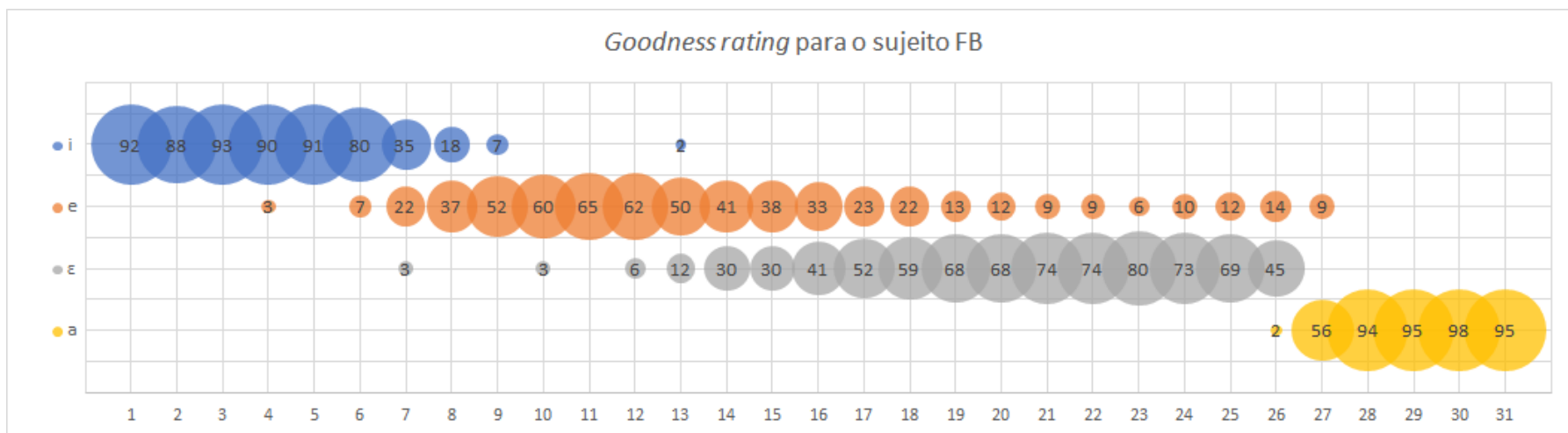
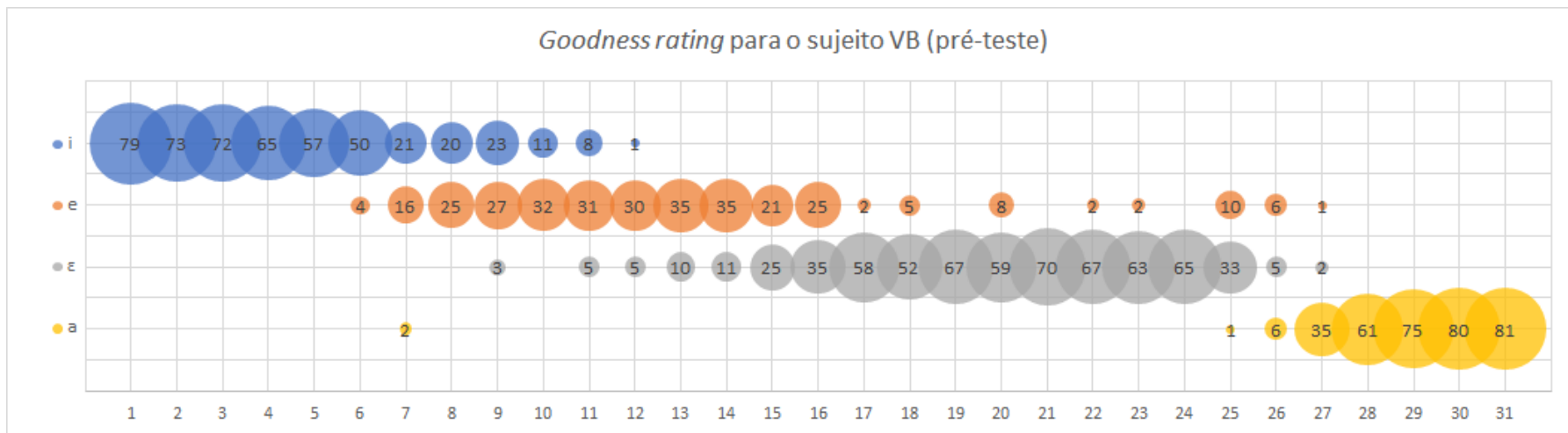
A curva de regressão logística para a transição da categoria /ε/ para não /ε/ no pré-teste (2/2)



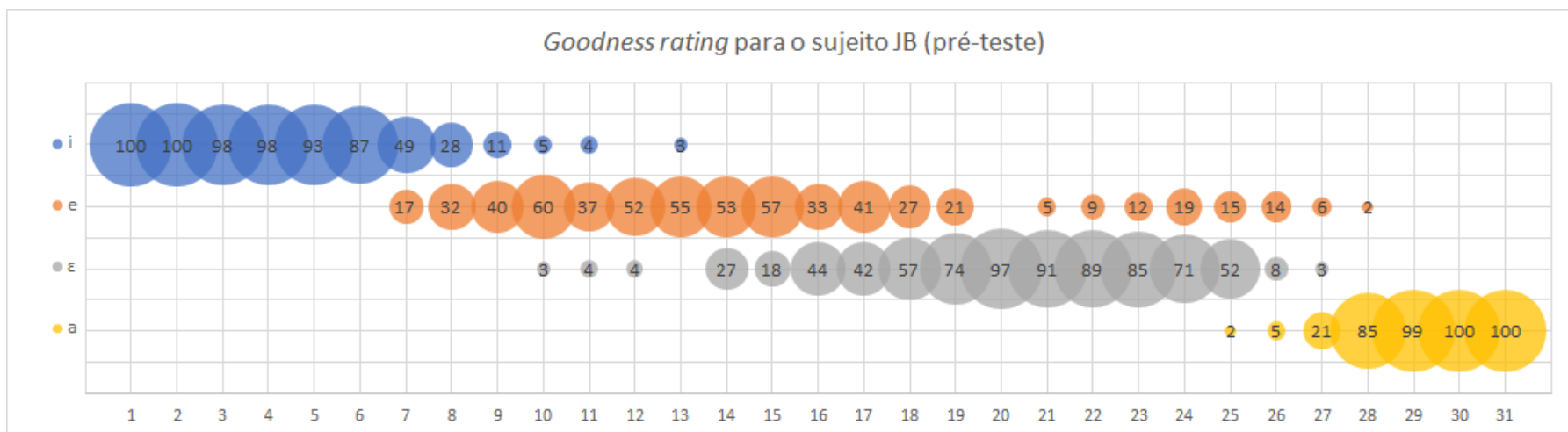
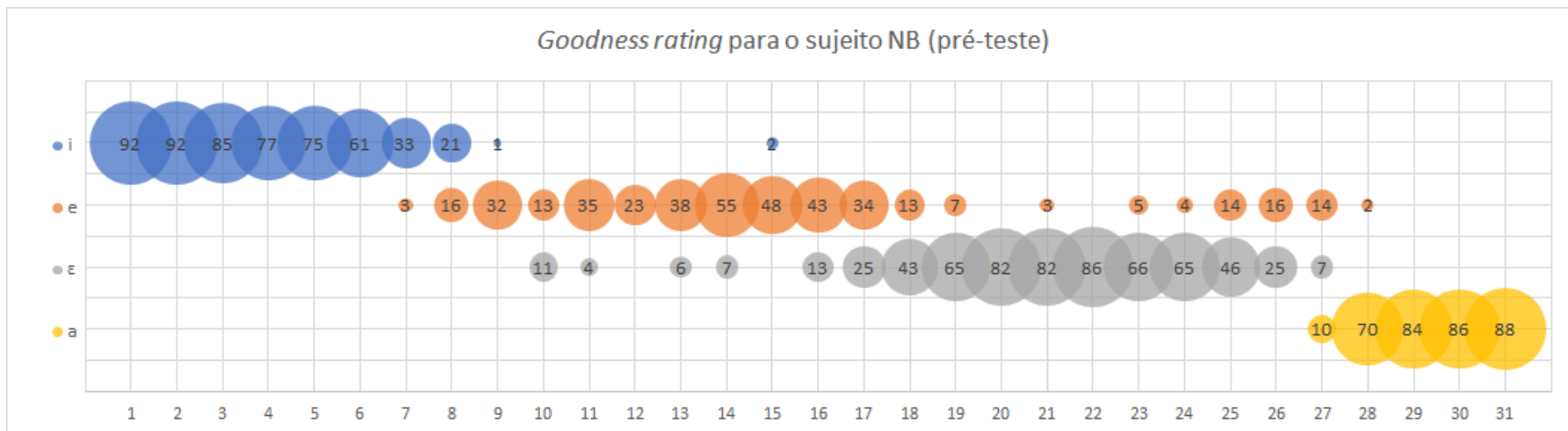
Goodness rating no pré-teste (1/4)



Goodness rating no pré-teste (2/4)



Goodness rating no pré-teste (3/4)



Goodness rating no pré-teste (4/4)

