

# Percepção e produção de vogais anteriores do inglês por falantes nativos de mandarim

Andréia Schurt Rauber\*

Anabela Rato e Ana Lúcia Silva \*\*

## Resumo

Este artigo apresenta os resultados de duas experiências que testaram a percepção e a produção de vogais anteriores inglesas por dez falantes nativos de mandarim. No que diz respeito à produção, foram medidos os dois primeiros formantes e a duração das vogais. Os resultados revelam que as informantes chinesas tiveram dificuldades em produzir vogais inglesas sem equivalentes na língua nativa (L1) (/ɪ, ε, æ/), mostrando tendência para perceber duas categorias da língua estrangeira (L2) como sendo apenas uma e para produzir as vogais da L2 com menor duração do que as informantes americanas que formavam o grupo de controlo. Quanto à percepção, os resultados de um teste de identificação mostram que a vogal da L2 com uma equivalente na L1 (/i/) foi a que teve taxas de identificação correcta mais elevadas. Os resultados demonstram que a performance perceptiva é superior à produtiva, e que a percepção e a produção estão relacionadas: as taxas de identificação mais elevadas no teste de percepção correspondiam a resultados melhores em termos de produção; e as taxas de identificação mais baixas a uma produção menos precisa.

**Palavras-chave:** vogais; percepção; produção; mandarim; inglês

\* Universidade Católica de Pelotas, Programa de Pós-Graduação em Letras, Pelotas, Brasil.  
asrauber@gmail.com

\*\* Universidade do Minho, Centro de Estudos Humanísticos, Braga, Portugal.  
asrato@ilch.uminho.pt; untuner@gmail.com

## Abstract

This paper reports the results of two experiments that tested the perception and production of English front vowels by ten native speakers of Mandarin. As regards production, the first two formants and duration were measured. The findings reveal that the Chinese participants had difficulty producing the English vowels with no L1 analogs (/ɪ, ε, æ/), tending to merge two L2 categories into one and to produce the L2 vowels with shorter duration than the Americans who formed a control group. As regards perception, the results of an identification test show that the L2 vowel with an L1 analog (/i/) was the one with highest identification rates. The results provide evidence that perception outperforms production, and that perception and production are related: higher identification rates in the perception test were related to better production results, lower identification rates were related to poorer production.

**Keywords:** vowels; perception; production; Mandarin; English

## 1. Introdução

Os aprendentes de L2 adultos caracterizam-se por ter um “sotaque estrangeiro perceptivo” (Strange, 1995: 22), que é o equivalente perceptivo de um sotaque estrangeiro falado. O modo como as categorias fonológicas da L2 são mapeadas em categorias preexistentes da L1, e o tipo de intervenção que poderia ser feita para atenuar falsas correspondências (treino perceptivo) têm atraído um número crescente de investigadores desde os anos 80.

Um dos investigadores mais influentes nesta área tem sido James E. Flege. No seu Modelo de Aprendizagem da Fala (*Speech Learning Model*—SLM), Flege (1995) coloca a hipótese de o estabelecimento completo de categorias fonéticas para a L1 impedir a subsequente formação de categorias para a L2. Segundo o autor, isto poderá acontecer porque a fonologia da L1 leva a que os estudantes da L2 filtrem diferenças acústicas perceptivas. Este obstáculo à formação de novas categorias deve-se ao mecanismo cognitivo da classificação por equivalência. Devido a esse mecanismo, os estudantes da L2 podem perceber uma categoria da L2 e uma categoria da L1 como sendo semelhantes o suficiente para serem consideradas equivalentes. Porém, a quantidade de exposição à L2 e/ou o treino na distinção de categorias da L1 e da L2 têm-se mostrado eficazes no aperfeiçoamento da percepção e da produção de sons da L2.

No entanto, o tempo de exposição não é em si suficiente. *Input* de boa qualidade, frequência de uso da língua, e idade de aprendizagem/aquisição são também variáveis que influenciam a aprendizagem/aquisição de sons de uma L2 (Flege, Bohn & Jang, 1997; Flege, MacKay & Meador, 1999). No caso dos estudantes que aprendem a L2 num contexto formal, a prática da L2 tende a circunscrever-se à sala de aula, e a exposição a *input* de elevada qualidade pode ser mínima (limitando-se à televisão, rádio ou a materiais didácticos, tais como CDs ou DVDs). Além disso, os estudantes tendem a interagir sobretudo com outros estudantes com sotaques estrangeiros. Assim, aulas da L2 a partir de uma idade pré-pubertária são aconselháveis, mas a filosofia “quanto mais cedo melhor” no que concerne aprender uma pronúncia da L2 que se aproxime da nativa nem sempre garante que este objectivo seja atingido.

Um outro aspecto que favorece a mestria de alguns sons e não de outros é a importância da detecção de diferenças entre palavras da L2 com contrastes mínimos, tal como foi dito por Best e Tyler (2007: 30) no seu Modelo de Assimilação Perceptivo (*Perceptual Assimilation Model-PAM-L2*). Isto é, se as palavras têm uma frequência de uso elevada ou se precisam de ser discriminadas para permitir que haja compreensão correcta (exemplo: *beach* ([bitʃ] – praia) e *bitch* ([bitʃ] – cadela)), a pressão para soar compreensível provavelmente encorajará os falantes da L2 a aprender a distinção.

Neste estudo, investigámos a forma como um grupo de falantes nativos de mandarim percebe quatro vogais anteriores inglesas (/i, ɪ, ε, æ/). As quatro vogais foram escolhidas tendo em consideração a teoria PAM-L2 dos contrastes mínimos que ocorrem em áreas fonológicas densas e em palavras com alta frequência de uso. A escolha de falantes de mandarim deve-se ao facto de, segundo Wang (1997: 16-17), o mandarim padrão conter um inventário vocálico de cinco fonemas (/i, y, e, a, u/) e nove alofones (/i, ɪ, ə, ʏ, o, ε, ɔ, æ, ɑ/). Assim, enquanto que o inglês tem os monotongos /i, ɪ, ε, æ/, o mandarim tem apenas as vogais anteriores /i/ e /e/ e a vogal central /a/ em posição tónica. Considerando o seu inventário da L1, é interessante investigar como falantes nativos de mandarim categorizam as vogais /i, ε, æ/ do inglês. Um breve resumo de estudos que investigaram a relação entre a percepção e a produção de sons será apresentado na próxima secção. A Secção 1.2 descreverá mais informação sobre o presente estudo.

### 1.1. Estudos prévios sobre percepção e produção de fala

O Modelo de Aprendizagem da Fala (SLM) desenvolvido por Flege (1995: 238) explica que “sem ‘alvos’ perceptivos correctos para guiar a aprendizagem sensora-motora dos sons da L2, a produção dos sons da L2 será incorrecta” [nossa tradução]<sup>1</sup>. No entanto, tal não significa que um indivíduo que perceba correctamente os sons da L2 seja capaz de os produzir de forma precisa. Há outros aspectos que podem dificultar a produção, tais como a restrição ou complexidade articulatória. Alguns estudos explicam a relação entre percepção e produção (por exemplo, Thomson, 2007; Bion, Escudero, Rauber & Baptista, 2006; Koerich, 2006; Flege, Bohn & Jang, 1997; Flege, MacKay & Meador, 1999; Rauber, Escudero, Bion & Baptista, 2005; Wang, 1997). Estes estudos revelaram que os sons que são percebidos de forma mais precisa são aqueles também produzidos de forma mais correcta. Destacam-se, como fundamentais para o presente estudo, os trabalhos de investigação de Thomson (2007) e Wang (1997), dado que ambos investigaram a percepção e a produção de vogais inglesas por falantes nativos de mandarim que adquiriram o inglês como segunda língua (ESL) e focaram um vasto conjunto de vogais, facilitando assim a compreensão da distribuição das vogais inglesas no espaço vocálico da L2 dos participantes.

Os resultados do estudo de Wang (1997) mostram que as vogais localizadas numa área densa do espaço vocálico da L2 (tal como as vogais anteriores inglesas) irão dificultar a produção e a percepção dos participantes chineses, cuja L1 possui poucas vogais nessa área. Os falantes nativos de inglês que avaliaram a inteligibilidade das produções dos informantes chineses tiveram dificuldade em distinguir as vogais dos pares /i-ɪ/ e /ɛ-æ/. Os resultados obtidos mostram que as vogais da L2 que têm equivalentes na L1 foram significativamente mais inteligíveis do que aquelas presentes apenas no inventário da L2. No que concerne a relação entre percepção e produção, os participantes da L2 obtiveram melhores resultados na percepção do que na produção das vogais /ɪ, ɛ, æ/ da L2 (ausentes no inventário da L1). Relativamente às vogais /i, e/ que têm equivalentes na L1, a produção superou a percepção.

<sup>1</sup> “Without accurate perceptual ‘targets’ to guide the sensorimotor learning of L2 sounds, production of the L2 sounds will be inaccurate.”

Thomson (2007) investigou o efeito do treino perceptivo na percepção e na produção das vogais inglesas por chineses. Os resultados do seu estudo mostram que houve uma melhoria significativa da produção após um período de três semanas de treino perceptivo, que incluiu um total de 4 sessões de treino e de 200 itens por sessão (20 repetições de cada uma das dez vogais alvo). As vogais alvo foram inseridas nos contextos /bV/ e /pV/ (V = vogal). Primeiro, as vogais foram apresentadas em pares e só depois contrastadas com vogais previamente trabalhadas. Foram treinados três grupos de participantes com diferentes estímulos: o Grupo 1 ouviu estímulos artificialmente manipulados de forma a produzir vogais longas, facilitando-lhes assim a detecção de informação acústica relevante; o Grupo 2 ouviu um subconjunto de vogais produzidas naturalmente (no contexto /CV/; C= consoante), que foram avaliadas por falantes nativos do inglês como sendo mais distantes da pronúncia mandarim (e mais próximas da pronúncia da L2) do que outras vogais produzidas pelo mesmo grupo de chineses; o Grupo 3 ouviu o subconjunto oposto de vogais inglesas naturais, as avaliadas como mais semelhantes às do mandarim. Os resultados corroboraram as observações de Wang (1997) relativamente às vogais anteriores /i/ e /e/, presentes no inventário da L1, que obtiveram os melhores resultados na identificação, enquanto que /ɪ, ε, æ/ foram frequentemente confundidas com outras vogais. Igualmente, as vogais com maiores percentagens de identificação foram também as produzidas de forma mais correcta. O treino aperfeiçoou as competências de percepção e de produção e o grupo treinado com as vogais inglesas consideradas as menos próximas de qualquer categoria do mandarim (ou seja, mais parecidas com a língua alvo) obteve os melhores resultados.

Considerando os resultados obtidos nos estudos supramencionados, o presente trabalho teve como objectivo aprofundar a investigação sobre a relação entre percepção e produção das vogais inglesas. O grupo de participantes é constituído por falantes nativos de mandarim que estudaram formalmente a produção de vogais, mas não fizeram nenhum treino perceptivo sistemático. A secção seguinte facultará informação mais detalhada sobre este estudo.

## 1.2 Justificação do presente estudo

Nos últimos dez anos, tem-se verificado, em Portugal, um aumento significativo do número de imigrantes chineses (de 2.782 imigrantes legalizados em 1999 para 13.331 em 2008 – SEF, 2009). Com excepção do projecto *Diversidade Linguística na Escola Portuguesa*, desenvolvido pelo *Instituto de Linguística Teórica e Computacional* (ILTEC), não têm sido desenvolvidos muitos estudos que investiguem as dificuldades com que os imigrantes chineses se deparam na aprendizagem/aquisição da língua portuguesa. O conhecimento sobre as dificuldades que os imigrantes têm para pronunciarem os sons do português europeu (PE) é útil para a preparação de materiais didácticos apropriados e para informar os professores dos ensinos primário e secundário sobre as melhores formas de lidar com os problemas de comunicação dos alunos. Considerando estes aspectos, foi reunido um grupo de dez estudantes universitários, que estudavam português na China, mas, durante oito meses, frequentaram o terceiro ano da sua licenciatura na Universidade do Minho, em Braga. Uma vez que os alunos estavam a estudar português e inglês, a percepção e a produção de sons específicos das duas línguas foram testadas e analisadas. Neste trabalho serão descritos os resultados do estudo sobre as vogais inglesas.

Partindo dos resultados de Wang (1997) e Thomson (2007), o estudo teve como objectivo testar as seguintes hipóteses:

- 1) os participantes terão dificuldade na produção das vogais /ɪ, ε, æ/ da L2, ausentes no inventário vocálico da L1: terão tendência para pronunciar como pertencentes a uma única categoria fonológica os pares de vogais /i/ e /ɪ/, e /ε/ e /æ/. No caso do primeiro par, a categoria /ɪ/ da L2 e a categoria /i/ da L1 serão percebidas como suficientemente semelhantes para serem consideradas equivalentes;
- 2) devido à ausência de equivalentes na L1, as vogais /ɪ, ε, æ/ serão as mais facilmente confundidas com outras (exemplo, /i/ confundida com /ɪ/, /ε/ com /æ/, /æ/ com /ʌ/) num teste de identificação de vogais.

## 2. Experiência A: Produção da fala

### 2.1 Informantes

O grupo observado foi constituído por dez informantes chinesas, com idades compreendidas entre os 19 e os 22 anos (média = 20,6 anos, desvio padrão (DP) = 0,9 anos). Este grupo incluía estudantes universitárias que tiveram aulas de inglês durante, pelo menos, oito anos, (média = 9,1 anos, DP = 1,8 anos) na China. Todas declararam que a sua pronúncia se aproximava a do inglês americano e afirmaram nunca terem tido aulas com um falante nativo de inglês. As informantes faziam parte de um grupo de 15 alunos de licenciatura chineses que estudaram português e inglês durante oito meses na Universidade do Minho através de uma parceria entre o Instituto Confúcio, sediado na Universidade do Minho, e a Universidade de Línguas Estrangeiras de Tianjin, China. Estes alunos são futuros professores de português na China, mas a maioria informou que, devido às dificuldades inerentes às condições de trabalho, poderão tornar-se professores de inglês e/ou tradutores de português/inglês-chinês. Assim, foi interessante observar o seu desenvolvimento fonológico em ambas as línguas, português e inglês, sendo que este artigo analisa apenas os dados do inglês.

Foi também seleccionado um grupo de controlo constituído por sete falantes nativos de inglês americano. As informantes eram mulheres com idades entre os 22 e os 39 anos (média = 32,3 anos, DP = 6,8 anos). Estas incluíam estudantes, professoras ou funcionárias administrativas da Universidade de St. Ambrose, em Davenport, EUA. Nenhuma destas informantes tinha formação em linguística e foram seleccionadas consoante a sua disponibilidade. Apesar de haver uma grande variação na produção de vogais nos Estados Unidos, as produções das informantes desta região específica do país foram seleccionadas aleatoriamente apenas para mostrar como é configurado o espaço vocálico dos falantes nativos de inglês quando o mesmo conjunto de palavras lido por um grupo de informantes de L2 é também lido por um grupo de falantes nativos. A preferência pelo inglês americano foi devida ao facto de o grupo de informantes chinesas ter sido exposto predominantemente à variante americana na China e em Portugal.

## 2.2 Procedimentos

A sessão de recolha de dados foi realizada oito meses após a chegada das informantes chinesas a Portugal. Durante esse período, elas tiveram alguma formação sobre a fonologia do inglês, mas não foi realizado nenhum treino sistemático de percepção ou de produção. A professora de inglês, que os acompanhou durante esse período, foi informando os alunos sobre o seu desempenho e corrigindo a sua produção durante as aulas, embora não o tenha feito de forma sistemática e nem a todos os alunos. Assim, apesar de os alunos não terem recebido o treino fonético ideal, a professora tentou alertá-los para a produção correcta das vogais inglesas.

Os instrumentos e procedimentos de teste foram os mesmos para os grupos de controlo e observado (chinesas). Primeiro, pediu-se aos participantes que lessem a frase veículo “*Say X now*” que integrava palavras com a estrutura silábica CVC contendo os monotongos ingleses (/i, ɪ, ε, æ/). Todas as palavras utilizadas neste estudo foram seleccionadas de acordo com o estudo de Hagiwara (1997) e continham uma vogal precedida por uma consoante oclusiva ou /h/, e seguida de uma consoante oclusiva (ver Tabela 1). O controlo do contexto consonântico facilitou a medição da duração, uma vez que as consoantes seleccionadas permitiram uma visualização mais perceptível dos pontos de início e de fim de cada vogal no oscilograma e espectrograma. Para minimizar os efeitos da influência da ortografia, antes de apresentar a frase a ser lida, era exibida uma imagem ilustrativa de uma palavra que continha a vogal alvo e dizia-se aos participantes que a palavra que a imagem representava deveria rimar com a palavra alvo da frase. Quando uma palavra não era lida de acordo com o esperado, principalmente devido à falta de atenção, a imagem era mostrada novamente e os participantes reliam a frase. As imagens, criteriosamente seleccionadas, representavam palavras simples de uso frequente para facilitar a sua associação à produção ideal da palavra alvo. As imagens adequadas para cada vogal foram /i/: *bee* (abelha), /ɪ/: *pig* (porco), /ε/: *egg* (ovo), e /æ/: *cat* (gato). Os conjuntos de imagem+frase foram apresentados de forma aleatória em diapositivos num computador portátil. Desta forma, os participantes viam, por exemplo, a imagem de um gato num diapositivo, pronunciavam a palavra isolada “*cat*” e depois liam a frase “*Say bat now*”,

mostrada no diapositivo seguinte. Foi pedido aos participantes que lessem as frases com entoação descendente e a ritmo normal. Para evitar a entoação característica de leitura de lista, incluiu-se apenas uma frase por diapositivo. A gravação dos dados do grupo de chinesas foi realizada numa sala silenciosa na Universidade do Minho e a do grupo de controlo numa sala silenciosa na Universidade de St. Ambrose, com um gravador Edirol R-09HR e um microfone unidireccional Edirol CS-15, a uma taxa de amostragem de 22kHz e quantização de 16-bit.

**Tabela 1.** Palavras utilizadas para incitar a produção das vogais alvo.

	/i/	/ɪ/	/ɛ/	/æ/
b_t	beat	bit	bet	bat
t_k	teak	tick	tech	tack
h_d	heed	hid	head	had

Cada palavra alvo foi lida três vezes pelo grupo observado e duas vezes pelo grupo de controlo. O número total de vogais produzidas pelo grupo observado foi de 360 (4 vogais x 3 contextos x 3 repetições x 10 participantes), e o número total de vogais produzidas pelo grupo de controlo foi de 168 (4 vogais x 3 contextos x 2 repetições x 7 participantes).

### 2.3 Medições das vogais

As vogais foram segmentadas e etiquetadas manualmente e posteriormente normalizadas, de acordo com a intensidade de pico, usando o programa Praat Versão 5.0.21 (Boersma & Weenink, 2008). Para medir a duração, o primeiro pico positivo da porção periódica foi considerado o início da vogal e o ponto onde a amplitude e complexidade da forma de onda diminuía foi considerado o seu fim. Os pontos de início e de fim da vogal foram marcados em cruzamentos zero (onde a forma de onda cruza a amplitude zero). As frequências para os dois primeiros formantes<sup>2</sup> (F1 e F2) foram medidas usando a análise de Codificação de Predição Linear (LPC), com cinco formantes por quadro. Os valores

<sup>2</sup> Dependendo da configuração do tracto vocal (por exemplo, grau de abertura da mandíbula, protusão labial, altura da língua), as suas frequências naturais de ressonância formam diferen-

medianos da porção equivalente aos 40% centrais de cada vogal foram calculados automaticamente e, se um valor medido fosse substancialmente diferente de outros encontrados para a mesma vogal produzida pelo mesmo participante, a vogal era medida novamente, mas desta vez manualmente, movendo-se o cursor para uma porção estável adjacente e recalculando-se os valores formânticos. Através deste procedimento, obtiveram-se valores em maior conformidade com o valor esperado. A mediana foi utilizada em detrimento da média para reduzir o efeito de erros de medição dos formantes, já que a mediana considera os valores do meio de uma distribuição. A distância Euclidiana foi calculada para os pares de vogais /i, ɪ/ e /ɛ, æ/ para verificar o quão distante uma vogal se encontrava de outra em termos de valores de F1 e F2. Assim, o objetivo foi comparar diferenças em termos de *distância* entre as vogais de cada par produzidas pelas chinesas e americanas e não comparar valores de formantes para cada vogal isolada.

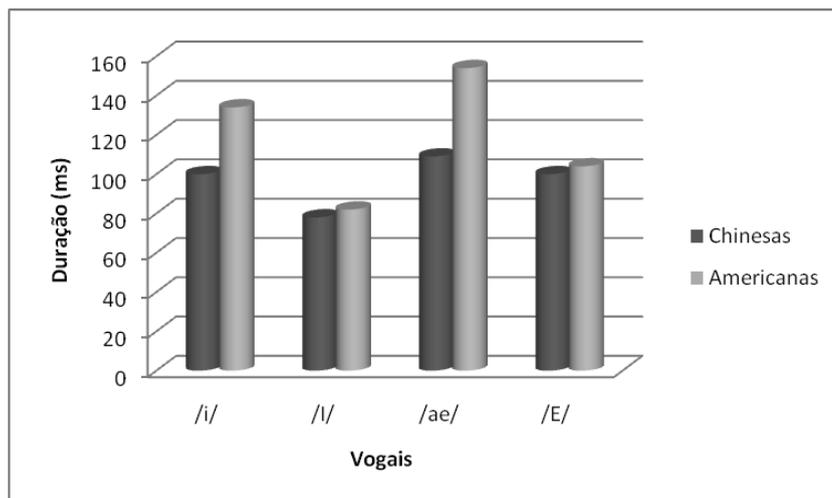
## 2.4 Resultados

Foram analisadas diferenças de performance na L2, comparando-se a produção do grupo de informantes chinesas à do de americanas. Os testes de normalidade Kolmogorov-Smirnov revelaram que os dados sobre duração e distância Euclidiana dos dois grupos apresentaram distribuição normal. Relativamente à duração, as diferenças entre os dois grupos ocorreram apenas para as vogais /æ/ e /i/ (/æ/:  $t(15) = -3,61$ ,  $p = 0,003$ ; /i/:  $t(15) = -3,40$ ,  $p = 0,004$ ), isto é, as vogais foram produzidas com duração significativamente menor pelas chinesas do que pelas americanas. Quanto à variação intra-grupo dos valores de duração entre as vogais produzidas pelas chinesas, uma análise de variância unifactorial (ANOVA) revelou diferenças significativas ( $F(3)=3,55$ ,  $p = 0,024$ ). Os testes *post-hoc* de Scheffé mostraram que apenas as vogais /æ/ e /ɪ/ diferiram significativamente uma da outra ( $p = 0,033$ ), sendo esta última a mais curta.

tes picos vocálicos que podem ser visualizados num espectro. Estas frequências de ressonância naturais do tracto vocal são chamadas de formantes (Johnson, 2003: 96).

**Tabela 2.** Valores médios de duração (em milissegundos), F1, F2 e distância Euclidian (DE, em Hz) dos grupos de chinesas e americanas (DP = desvio padrão)

		/i/	/ɪ/	/æ/	/ɛ/	DE /i-ɪ/	DE /æ-ɛ/
<b>Chinesas</b>							
	Dur	100 (23,3)	78 (13,8)	109 (26,7)	100 (24,0)		
Média (DP)	F1	421 (32,1)	438 (36,7)	823 (93,7)	753 (80,3)	115 (121)	104 (34)
	F2	2754 (143,9)	2648 (184,5)	2077 (180,4)	2098 (177,4)		
<b>Americanas</b>							
	Dur	134 (14,4)	82 (16,4)	154 (23,3)	104 (18,0)	580 (81)	129 (42)
Média (DP)	F1	393 (29,4)	565 (32,1)	816 (27,8)	713 (15,7)		
	F2	2744 (146,5)	2228 (154,5)	1998 (97,1)	1968 (79,1)		



**Figura 1.** Valores médios de duração (em milissegundos) de vogais produzidas pelas informantes chinesas e americanas

Os resultados sobre a qualidade espectral revelaram que as informantes chinesas tenderam a agrupar as vogais dos pares /i-ɪ/ e /æ-ɛ/ (ver Figura 2). Relativamente à distância Euclidiana entre as vogais do contraste /i-ɪ/, os testes-t de amostras independentes mostraram diferenças significativas entre as chinesas e as americanas ( $t(15) = -8,82, p < 0,001$ ). Todavia, para o contraste vocálico /æ-ɛ/, não foi encontrada nenhuma diferença significativa. No entanto, no caso deste contraste em específico, apesar de não se verificar uma diferença significativa nos valores de distância Euclidiana, as duas vogais produzidas pelas chinesas não seguiram a mesma direcção das produzidas pelas americanas: a vogal /ɛ/ ou tem valores de F2 muito semelhantes à /æ/ ou é mais anteriorizada, diferentemente das americanas, cuja vogal /æ/ é mais centralizada do que /ɛ/. Para além deste aspecto, as duas vogais foram produzidas de forma mais anteriorizada pelas chinesas do que pelas americanas.

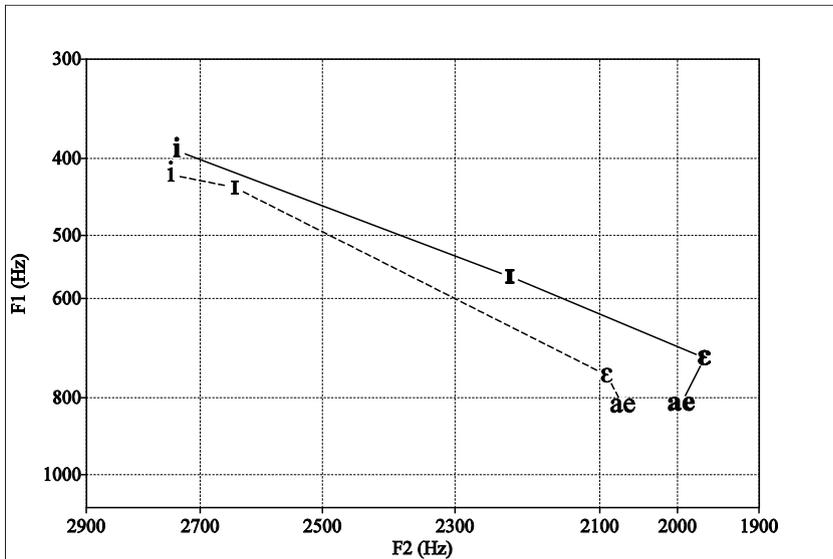


Figura 2. Vogais anteriores do inglês produzidas pelas americanas (linha) e chinesas (pontilhado)

### 3. Experiência B: Percepção de fala

#### 3.1 Informantes

As mesmas falantes nativas de mandarim descritas na Secção 2.1 formaram o grupo observado na Experiência B. Devido à disponibilidade de informantes, o grupo de controlo foi formado por cinco americanos, um homem e quatro mulheres, com idades entre 22 e 24 anos (média = 22,5 anos, DP = 0,9 anos), alunos de mestrado do curso de Terapia Ocupacional da Universidade de St. Ambrose, em Davenport. Todos os informantes afirmaram não ter problemas de audição.

#### 3.2 Estímulos

As vogais /i, ɪ, ε, æ, ʌ/ foram inseridas em estruturas CVC(C) (ver Tabela 3) e gravadas por quatro estudantes universitários de Sacramento e Berkeley, Califórnia. Os informantes eram dois homens e duas mulheres, com idades entre 25 e 27 anos (média = 25,8 anos, DP = 0,9 anos). As palavras reais e artificiais foram inseridas na frase veículo “X and Y sound like Z” (“X e Y soam como Z”), lidas com entoação decrescente e a ritmo normal. As três palavras alvo da frase veículo continham a mesma vogal, por exemplo: “*Beat and Pete sound like seat.*” (*Beat* e *Pete* soam como *seat*). Os procedimentos e os informantes foram os mesmos de Rauber (2006). A vogal central /ʌ/ foi incluída para testar se os informantes falantes da L2 confundiriam esta vogal com outras vogais baixas.

**Tabela 3.** Estímulos utilizados no teste de percepção

/i/	/ɪ/	/ε/	/æ/	/ʌ/
beat	bit	bet	bat	but
keep	kit	kept	cat	cut
Pete	Pitt	pet	pat	putt
seat	sit	set	sat	shut
teak	tick	tech	tack	tuck
teat	tit	tet	tat	tut

Cada palavra foi apresentada isoladamente quatro vezes, tendo sido duas vezes pronunciada por uma mulher e duas por um homem. Assim, as informantes chinesas foram expostas a diferentes vozes, o que fez com que tivessem que ignorar diferenças audíveis (qualidade da voz), mas foneticamente irrelevantes em termos de variação intra-categórica. O número total de palavras ouvidas foi de 120 (5 vogais x 6 palavras x 4 repetições). As palavras alvo foram segmentadas e normalizadas de acordo com o pico de intensidade no software Praat e depois apresentadas isoladamente às informantes através de um teste de identificação, também elaborado e apresentado no Praat. Mais detalhes sobre o teste são descritos nas subsecções seguintes.

### 3.3 Procedimento

O teste foi realizado nos computadores de um laboratório de línguas e durou aproximadamente 30 minutos. Os grupos observado e de controlo ouviram os estímulos através de auriculares e tiveram que seleccionar o símbolo fonético que melhor representasse o som que ouviam. Já que nem todos estavam familiarizados com os símbolos fonéticos, as palavras *beat*, *bit*, *bet*, *bat* e *but* foram utilizadas como as etiquetas dos botões apresentados no ecrã do computador através do software Praat. Antes do teste, os informantes leram as palavras para se familiarizarem com as opções. Foi realizada uma sessão de treino com 10 estímulos, usando os próprios estímulos do teste. As palavras eram automaticamente aleatorizadas cada vez que o teste era iniciado. O grupo de controlo desta experiência foi importante para validar os estímulos. Estes informantes foram testados numa sala silenciosa utilizando um computador portátil e auriculares. Os informantes americanos identificaram todos os estímulos /i, ɪ/ correctamente. Contudo, os resultados médios de identificação foram de 98,3% para /ɛ/ e 96,7% para /æ/. Assim, os estímulos foram considerados fiáveis para testar o grupo de chinesas. Os testes de percepção e de produção foram realizados em dias diferentes.

### 3.4 Resultados

Tal como referido na Secção 2.4, as informantes chinesas tenderam a não produzir as diferenças em termos de duração e de qualidade espectral

entre os contrastes /i/-/ɪ/ e /ɛ/-/æ/ como as americanas. Uma explicação possível para este resultado é a dificuldade em *perceber* as vogais da L2 que não estão presentes no inventário da L1 e assimilar duas vogais como pertencentes a uma única categoria. Esta dificuldade pode ser observada na Tabela 4, que mostra os resultados individuais e as médias de identificação do grupo de chinesas.

**Tabela 4.** Resultados individuais e médios das informantes chinesas

Part.	/i/	/ɪ/	/æ/	/ɛ/
1	95,8	83,3	83,3	37,5
2	95,8	91,7	100,0	50,0
3	75,0	41,7	62,5	25,0
4	91,7	95,8	79,2	54,2
5	100,0	87,5	62,5	70,8
6	95,8	91,7	50,0	83,3
7	100,0	100,0	75,0	62,5
8	87,5	54,2	91,7	50,0
9	62,5	75,0	70,8	62,5
10	100,0	91,7	79,2	70,8
Média (DP)	75,4 (14,8)	56,7 (17,1)	90,4 (12,4)	81,3 (19,1)

Uma vez que os dados de percepção não apresentaram uma distribuição normal, realizámos os testes não-paramétricos de Wilcoxon. Uma matriz de confusão (ver Tabela 5) mostra que a vogal /i/, presente no inventário da L1 (mandarim), foi a vogal que obteve as maiores percentagens de identificação correcta, sendo apenas ocasionalmente confundida com /ɪ/. A vogal /ɪ/ teve a tendência para ser confundida com /ɛ/ em vez de /i/. Quando as informantes ouviam /æ/, muitas identificavam esta vogal como sendo /ʌ/ mais frequentemente do que como sendo /ɛ/, provavelmente porque tanto /æ/ como /ʌ/ estão mais próximas do seu /a/ do que do seu /e/ da L1, ou seja, a referência era uma vogal baixa e central. Finalmente, quando /ɛ/ era equivocadamente identificado, a confusão tendeu a ser com /æ/.

**Tabela 5.** Percentagens médias de identificação dos estímulos do inglês por vogal. Os valores representam a percentagem das vogais reconhecidas como pertencentes a cada categoria vocálica do inglês.

Estímulo	Vogal identificada				
	/i/	/ɪ/	/æ/	/ɛ/	/ʌ/
/i/	90,4	9,6	—	—	—
/ɪ/	5,0	81,3	—	13,7	—
/æ/	—	—	75,4	9,2	15,4
/ɛ/	—	3,3	39,2	56,7	0,8

#### 4. Discussão e notas finais

Este estudo investigou a percepção e a produção de vogais anteriores do inglês por 10 falantes nativas de mandarim que receberam alguma instrução sobre produção durante um curso de nível pós-intermédio de inglês. Os resultados confirmam as hipóteses do estudo. Quanto aos resultados de produção, as informantes tiveram dificuldade em produzir as vogais da L2 /ɪ, ɛ, æ/, que são vogais sem equivalentes na L1. Relativamente à qualidade espectral das vogais altas, as chinesas produziram valores de F1 e F2 para /i/ muito próximos dos das americanas, mas a distância Euclidiana entre as duas vogais altas foi muito pequena. Por outro lado, para o contraste /æ-ɛ/, os valores formânticos mostram que as informantes chinesas produziram a vogal /æ/ mais baixa do que /ɛ/, mas esta foi produzida de forma tão centralizada como aquela. Como previsto pelo SLM de Flege (1995), a tendência verificada foi a de agrupar duas categorias como pertencentes a uma única categoria, como foi o caso das duas vogais altas (/i/ e /ɪ/) e das duas vogais baixas (/ɛ/ e /æ/). Em termos de duração, os resultados revelam que as informantes chinesas pronunciaram as vogais frouxas /ɪ/ e /ɛ/ com valores de duração similares aos das americanas. Embora as vogais tensas /i/ e /æ/ tenham sido produzidas com maior duração do que as frouxas, foram significativamente mais curtas do que as vogais alvo. Isto indica que as falantes de L2 *fazem* uso de informação duracional para distinguir entre vogais tensas e frouxas, mas não tanto como as americanas.

No que concerne a percepção, a vogal da L2 com uma equivalente na L1 (/i/) foi a que teve maiores índices de identificação (90,4%). A segunda vogal mais bem percebida foi /ɪ/ (81,3%), seguida por /æ/ (75,4%). A vogal frouxa /ɛ/ teve os menores índices de identificação (56,7%), sendo muito frequentemente identificada como /æ/ (39,2%). Os resultados para /ɪ/ e /æ/ corroboram estudos como os de Bion *et al.* (2006), Flege (1995), Flege *et al.* (1997, 1999) e Rauber *et al.* (2005) ao indicar que a percepção precede a produção e os resultados gerais indicam que há uma relação entre altas percentagens de identificação no teste de percepção e melhores resultados de produção (/i/), e entre baixas percentagens de identificação e uma produção mais deficiente (/ɪ, ɛ, æ/). Uma razão para essa tendência pode ser o facto de que alguns aspectos da percepção mudam antes do que as alterações correspondentes sejam implementadas na produção; no entanto, não se pode descartar a possibilidade de que mudanças na produção aconteçam primeiro ou que ocorram mesmo que não haja mudanças na percepção.

A dificuldade em perceber e produzir categorias da L2 inexistentes na L1 leva a uma questão polémica em relação ao momento mais propício para fornecer treino fonético ou instrução sobre pronúncia a alunos de uma L2. Se um aluno precisa de aprender, por exemplo, léxico e regras gramaticais desde as fases mais iniciais de aprendizagem desta L2, este também deveria ser instruído quanto às diferenças entre a estrutura sonora da L1 e da L2. É comum desprezar este tipo de informação linguística em contextos formais de aprendizagem. A opção de soar mais ou menos inteligível é certamente do aluno, mas incluir comentários e/ou instrução para direccionar o foco de atenção e consciencializá-los do funcionamento fonético-fonológico da L2 não podem ser vistas como não sendo importantes. As informantes chinesas deste estudo tiveram uma média de oito anos de instrução sobre a língua inglesa, mas nunca tinham recebido treino ou instrução específica sobre a pronúncia da L2. Portanto, o princípio de “o quanto antes, melhor” para adquirir uma boa pronúncia, tão comumente encontrado nas teorias sobre idade de aquisição/aprendizagem, deveria também ser aplicado ao início da exposição formal à L2, independentemente da idade: quanto mais cedo houver consciência sobre as regras fonológicas da L2, mais facilmente se evitam problemas de inteligibilidade relacionados com a produção.

## Agradecimentos

Gostaríamos de agradecer às informantes chinesas que se voluntariaram para este estudo; aos professores, alunos e funcionários da Universidade St. Ambrose, em Davenport-Iowa, por concordarem em fazer parte do grupo de controlo das duas experiências; ao CEHUM pelo financiamento parcial da viagem para a recolha de dados nos Estados Unidos; e aos valiosos comentários dos pareceristas anónimos.

## Referências

- BEST, C. T., & TYLER, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn, & M. J. Munro (Eds.), *Language Experience in Second language Speech Learning. In honor of James Emil Flege* (pp. 13-34). Amsterdam: John Benjamins.
- BION, R. A. H., ESCUDERO, P., RAUBER, A. S., BAPTISTA, B. O. (2006). Category formation and the role of spectral quality in the perception and production of English front vowels. *Proceedings of INTERSPEECH'2006 ICSLP* (pp. 1363-1366). Pittsburgh, USA.
- BOERSMA, P. & WEENINK, D. (2008). *Praat: Doing phonetics by computer* (Version 5.0.21) [Computer program]. Disponível em <http://www.praat.org> [consultado em 22 de Abril de 2008].
- CRYSTAL, D. (1997). *English as a global language*. Cambridge: Cambridge University Press.
- FLEGE, J. E. (1995). Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.
- FLEGE, J. E., BOHN, O.-S., & JANG, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25, 437-470.
- FLEGE, J. E., MACKAY, I. R. A., & MEADOR, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America*, 105(5), 2973-2987.
- JENKINS, J. (2009). *World Englishes: A resource book for students* (2. ed). Routledge.
- JOHNSON, K. (2003). *Acoustic and auditory phonetics* (2 ed.). Malden/Oxford/Victoria: Blackwell.
- KOERICH, R. D. (2006). Perception and production of vowel paragoes by Brazilian EFL students. In B. O. Baptista, & M. A. Watkins (Eds.), *English with*

- a Latin Beat: Studies in Portuguese/Spanish – English Interphonology* (pp. 91-104). Amsterdam: John Benjamins.
- HAGIWARA, R. (1997). Dialect variation and formant frequency: The American English vowels revisited. *Journal of the Acoustical Society of America*, 102 (1), 655-658.
- ILTEC – Instituto de Linguística Teórica e Computacional (2009). *Diversidade linguística na escola portuguesa*. Disponível em <http://www.iltec.pt/divling/index.html> [consultado em 20 de Setembro de 2009].
- RAUBER, A. S. (2006). *Perception and production of English vowels by Brazilian EFL speakers*. Unpublished doctoral dissertation. Florianópolis: Universidade Federal de Santa Catarina.
- RAUBER, A. S., ESCUDERO, P., BION, R. A. H., BAPTISTA, B. O. (2005). The interrelation between the perception and production of English vowels by native speakers of Brazilian Portuguese. *Proceedings of INTERSPEECH'2005 - EUROSPEECH* (pp. 2913-2916). Lisboa, Portugal.
- SEF – Serviço de Estrangeiros e Fronteiras (2009). *Estatísticas*. Disponível em <http://www.sef.pt/> [consultado em 20 de Setembro de 2009].
- STRANGE, W. (1995). Cross-language study of speech perception: A historical review. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 3-45). Timonium, MD: York Press.
- THOMSON, R. I. (2007). *Modeling L1/L2 interactions in the perception and production of English vowels by Mandarin L1 speakers: A training study*. Unpublished PhD dissertation. Edmonton: University of Alberta.
- WANG, X. (1997). *The acquisition of English vowels by Mandarin ESL learners: A study of perception and production*. Unpublished master's thesis. Vancouver: Simon Fraser University.

