

# TESTES ORAIS INDIVIDUAIS E EM PARES: UM ESTUDO DAS DIFERENÇAS DE DESEMPENHO

(Individual and Pair Speaking Test Formats: a Study  
of Differences in Performance)

Thaís Barbosa MAROCHI  
(FACEAR – Faculdade Educacional de Araucária (PR).  
Doutoranda – Universidade Federal do Paraná)

**RESUMO:** *O objetivo desta pesquisa foi verificar se há diferença no desempenho caso os candidatos façam o teste oral sozinhos ou em pares. Duas versões do First Certificate in English foram aplicadas para dez alunos em duas situações diferentes (examinador-candidato e candidato-candidato-examinador). Todos os testes foram gravados, transcritos e analisados estatisticamente. Após o segundo teste, os participantes responderam a um questionário para informar suas opiniões sobre os testes. Apesar de haver uma tendência em falar mais, usar uma maior diversidade de vocabulário e uma maior complexidade gramatical quando em pares, oito dos dez candidatos manifestaram uma preferência pela realização do teste individual.*

**PALAVRAS-CHAVE:** *avaliação; língua inglesa; proficiência; oralidade.*

**ABSTRACT:** *This research is aimed at checking whether there are differences or not on the candidates' test performances in case the test is taken individually or in pairs. Two versions of the First Certificate in English were applied on ten students in two different occasions (examiner-candidate and candidate-candidate-examiner). All tests were recorded, transcribed and statistically analyzed. After the second test, the participants answered a questionnaire to give their opinion about the tests. Although there is a bias to speak more, use a more diverse vocabulary and when in pairs, eight out of ten candidates revealed they would rather take the test individually.*

**KEY-WORDS:** *language testing; English language; proficiency; speaking.*

## 1. Introdução

Os testes orais de proficiência em língua inglesa da *University of Cambridge Local Examinations Syndicate* (UCLES) são elaborados para serem utilizados com dois (excepcionalmente três) candidatos e dois examinadores (um interlocutor e um avaliador). Ao interlocutor cabem duas tarefas: promover a interação entre os candidatos – por meio de tarefas previamente especificadas – e fornecer uma avaliação global para o desempenho de cada candidato, enquanto o avaliador usa uma escala analítica.

Muitos fatores estão envolvidos nesse tipo de avaliação, tais como os candidatos, os examinadores, o critério de avaliação, as tarefas. Dentre esses fatores, o que nos chamou mais a atenção foi com relação à participação dos próprios candidatos, ou seja, dois candidatos são escolhidos para, juntos, fazerem o teste. Apesar de o teste em pares ser também adotado por muitas universidades, institutos de línguas e outros cursos, há muitas controvérsias se essa situação melhor favorece o desempenho. Os estudos de Ross e Berwick (1992) e Young e Milanovic (1992) mostram que o discurso é muito assimétrico em um teste individual, o que não ocorre em um teste em pares. Együd e Glover (2001), tendo por base uma pesquisa com estudantes húngaros, concluem que os alunos ficam mais satisfeitos e produzem um discurso de melhor qualidade quando interagem com outro aluno. Por outro lado, Foot (1999:39) critica o teste em pares, mencionando que a diferença de idade, interesses e competência lingüística entre dois candidatos podem causar problemas quando da realização de um teste oral. Segundo Foot (1999), alguns estudantes não gostam do teste em pares e, quando podem escolher, preferem fazê-lo individualmente. Essa preferência é também observada por professores da Universidade Federal do Paraná, onde a presente pesquisa foi realizada. Assim, nosso objetivo é verificar se os candidatos apresentam ou não o mesmo desempenho, caso façam o teste sozinhos.

Para tanto, inicialmente é discutido o tipo de teste utilizado e suas características. Logo após, é mostrada a pesquisa propriamente dita. Os métodos são demonstrados e os sujeitos descritos. Em seguida, é feita a análise dos dados e a discussão dos resultados e, finalmente, encerra-se com algumas considerações finais a respeito dos dados obtidos.

## 2. Os testes orais de proficiência

Dentre os diversos testes orais em LE (língua estrangeira) utilizados atualmente, há os testes de proficiência, os quais são baseados na especificação do que o candidato é capaz de fazer na língua-alvo para ser considerado proficiente. No entanto, a própria palavra “proficiente” causa problemas. Segundo Scaramucci (2000), existem basicamente duas maneiras de abordar o termo. Não há um conceito “absoluto” equivalente a tudo ou nada, ou seja, proficiente ou não proficiente. O que há é um conceito relativo, que procura levar em conta a especificidade da situação de uso futuro da língua. Nessa visão, existe uma gradação, desde o menos proficiente até proficiente nativo ou bilíngüe. Para o nosso caso, essa primeira abordagem é a mais adequada.

Já a segunda maneira de definir proficiência, segundo a autora, é contrastá-la com rendimento, ou seja, é específica e local, à medida que descreve a aprendizagem de determinado programa. Essa distinção, conforme mencionado por Scaramucci (2000:15), determina procedimentos diferentes na elaboração dos instrumentos de avaliação: o de rendimento é baseado no programa, materiais e currículo; o de proficiência tem um compromisso com o construto teórico e tem suas especificações baseadas nas necessidades do público alvo com relação ao uso futuro da língua, pelo menos em suas visões mais contemporâneas. Essa distinção, de acordo com a autora, tem sido fundamental para a elaboração e a escolha de exames mais apropriados, bem como para a validação deles.

Segundo Lazaraton (2002:116), os alunos normalmente procuram os exames de proficiência para melhorar suas perspectivas de emprego, para estudos posteriores, para se prepararem para viajar ou morar no exterior ou simplesmente porque querem um certificado reconhecido internacionalmente de seu nível alcançado em uma LE.

Todos os exames de inglês elaborados pela UCLES são constituídos de cinco partes (leitura, escrita, uso do inglês, compreensão auditiva e fala). Nosso interesse está na avaliação oral e, mais especificamente, no *First Certificate in English* (FCE), por ser um exame considerado válido, confiável, amplamente conhecido e porque, de acordo com Lazaraton (2002), é o mais procurado pelos estudantes de LE. Além disso, representa o terceiro nível de proficiência (entre cinco) da UCLES e está atrelado a um sistema

internacional de avaliação de línguas européias estabelecido pela *Association of Language Testers in Europe* (ALTE).

Para ser aprovado no FCE, os alunos devem possuir um nível pós-intermediário de proficiência em inglês como LE, ou seja, devem ser capazes de lidar com as principais estruturas da língua com alguma confiança, demonstrar conhecimento de uma variedade razoável de vocabulário e uso apropriado de estratégias comunicativas em situações sociais (*University of Cambridge Local Examinations Syndicate*, 2003:6). Devem ser usuários independentes, conforme mencionado pela ALTE e possuir um comando efetivo de uma língua em situações variadas (*University of Cambridge Local Examinations Syndicate*, 1999:3).

De acordo com a *University of Cambridge Local Examinations Syndicate* (2003: 46), o teste oral do FCE está dividido em quatro partes. Na primeira, cada candidato deve responder às perguntas do interlocutor e ouvir o que seu parceiro tem a dizer. Na segunda parte, são fornecidas gravuras e cada candidato deve comparar e contrastar as gravuras e falar por aproximadamente um minuto. Cada candidato deve mostrar habilidade de organizar seus pensamentos e idéias, expressar-se em linguagem apropriada e prestar atenção ao seu parceiro para poder comentar algo sobre o que foi falado. Na terceira parte, são fornecidas instruções orais e gravuras e cada candidato deve ser capaz de se interar na conversa com a finalidade de expressar opiniões e idéias. E finalmente, o interlocutor faz perguntas para encorajar os candidatos a ampliar e discutir mais os tópicos introduzidos na parte 3.

Para este estudo, utilizamos apenas as partes 3 e 4 do teste com o objetivo de termos apenas um critério de avaliação. Essas duas partes foram escolhidas por demonstrar a habilidade dos candidatos em participar ativamente de uma conversa (parte 3) e discutir mais sobre os tópicos por meio de perguntas feitas pelo interlocutor (parte 4).

### 3. Metodologia

A pesquisa foi realizada nas dependências da Universidade Federal do Paraná, na cidade de Curitiba, estado do Paraná. Todos os testes feitos foram gravados e tiveram uma duração aproximada de sete minutos. Os

dados foram coletados durante o mês de maio e início do mês de junho do ano de 2004 nas seguintes fases:

*1ª. fase*

O teste foi aplicado em duplas. Tivemos um total de seis duplas participantes, das quais a primeira foi utilizada apenas como experimentação.

*2ª. fase*

Uma versão diferente do teste\* foi aplicada aos alunos individualmente.

*3ª. fase*

Foi aplicado um questionário para os alunos imediatamente após a realização do teste individual, com duas finalidades: obtenção de dados pessoais e conhecimento da opinião sobre as duas diferentes situações de teste.

Logo em seguida, iniciamos a transcrição dos testes, tanto individuais como em duplas, como também a tabulação dos dados fornecidos pelo questionário.

Nosso objetivo foi verificar se há a variação da linguagem produzida pelos candidatos em duas diferentes situações de teste oral (candidato-candidato-examinador e candidato-examinador). Para isso, observamos na linguagem produzida pelos candidatos, nas duas diferentes situações:

- a quantidade de linguagem produzida em um determinado período de tempo;
- a complexidade gramatical utilizada;
- diversidade do vocabulário, em termos de número de palavras diferentes.

### *3.1. Sujeitos da pesquisa*

Contamos com a participação de doze alunos da Universidade Federal do Paraná. Duas alunas participaram apenas da primeira fase, pois devido

---

\* Entenda-se por uma versão alternativa do mesmo teste. Quando da realização do teste oral do FCE os examinadores recebem o material contendo várias versões que podem ser aplicadas aos candidatos. Essas versões são totalmente equivalentes em termos de dificuldade, formato etc.

à péssima gravação, grande parte do teste ficou incompreensível. Por ser o primeiro teste aplicado, foi utilizado apenas como uma experimentação.

Dentre os dez alunos restantes, havia três homens e sete mulheres, e uma examinadora, no caso, a própria pesquisadora. A idade dos alunos variava entre 19 e 48 anos e todos estavam cursando o último ano de Letras da Universidade Federal do Paraná. Dos dez alunos, oito cursaram o Ensino Médio em escolas públicas, e dois em escola particular; quatro alunos estudaram espanhol, sendo: dois de nível intermediário, um de nível básico e um de nível avançado. Quatro alunos freqüentaram somente as aulas de inglês oferecidas durante o Ensino Fundamental e o Ensino Médio e os demais freqüentaram institutos de línguas por pouco tempo (de 3 meses a 3 anos). Apenas uma aluna teve experiência no exterior, morou em Londres por quatro anos, no período compreendido entre abril/91 a abril/95. Todos os alunos fizeram as partes 3 e 4 do exame oral do FCE para fins específicos da pesquisa em questão. A partir de agora, esses alunos passam a ser denominados de sujeitos (S) com as letras iniciais do alfabeto: **Sa, Sb, Sc, Sd, Se, Sf, Sg, Sh, Si, Sj**.

### 3.2. *Os dados*

Para analisarmos a velocidade da fala, ou seja, para contarmos o número de palavras por minuto e a diversidade do vocabulário (em termos de número de lexemas diferentes), precisamos inicialmente verificar o que consideramos como palavra e lexema diferente.

#### 3.2.1. *A palavra*

Segundo Huddleston e Pullum (2002:44), a palavra é a menor unidade da sintaxe e o termo, como usado na gramática tradicional, tem dois sentidos diferentes: o sintático e o léxico. Assim se consideramos *paint* e *painting* sob o aspecto sintático, temos duas palavras distintas. Já sob o aspecto lexical, são duas formas diferentes do mesmo lexema. Vejamos um outro exemplo: *get up*. Sob o aspecto sintático temos duas palavras, porém sob o aspecto lexical, podemos considerar como um verbo composto ou complexo, ou seja, um só lexema (levantar).

Para a contagem do número total de palavras, usamos o aspecto sintático, por estarmos interessados na quantidade de linguagem produzida. Tivemos como base Lennon (1990) e Fortkamp (1998) e consideramos ainda:

a) As contrações:

- uso do *not*: de acordo com Radford (1997:232), o *not* pertence a uma categoria separada NEG (constituente negativo) e funciona como uma palavra sintaticamente negativa (NEGP). Se a posição NEG for ocupada por *n't* ao invés do *not*, há necessidade do *n't* estar ligado a um auxiliar e constitui uma forma complexa (auxiliar + *n't*). É o caso de *don't*, *doesn't*, *won't*. Essa forma negativa comporta-se como qualquer outro auxiliar. Se fizermos, por exemplo, a interrogativa-negativa, ocorre a inversão, ou seja, os auxiliares movem-se da direita para a esquerda do sujeito (do INFL para o COMP), comportando-se apenas como uma palavra. Ex: *He won't come.* / *Won't he come?* Porém, se transformarmos a mesma oração: *He will not come* em interrogativa, teremos: *Will he not come?*, ocorre o movimento do auxiliar *will* do INFL para o COMP e o *not* permanece na posição de ADV do VP. Assim, consideramos *won't* como sendo uma palavra.

- ocorrências como as de *it's*, *he'll*, foram consideradas duas palavras. Radford (1997:250) sugere que, nesse caso, ocorre puramente um processo fonológico e que *it* e *is* permanecem duas palavras separadas. É o que o autor chama de “movimento PF” (*movement of phonetic features*). Podemos ainda comprovar a existência de duas palavras com a inversão que ocorre nas interrogativas. Um exemplo citado por Radford (1997:218): *You'll marry me.* (afirmativa).

*Will you marry me?* (interrogativa \_ ocorre a inversão do sujeito e auxiliar, o movimento do I para o C).

- outros casos – Conforme Radford (1997:269), *want to* pode ser reduzido a *wanna* (cliticização de *to* em *want*, assimilação de /nt/ em /nn/ e a redução /nn/ em /n/, como no exemplo :*I want to go home.*

*I wanna go home.*

Algo semelhante ocorre com *gonna* (*going* + *to*), o qual é somente um caso de redução e foi considerado como duas palavras. Ou ainda

com *gotta* (*has/have* + *got to*), neste caso considerado como duas palavras.

- casos de 's \_ historicamente o 's é uma flexão como a que ocorre com o plural -s. Portanto, em uma frase como: *John's father*, temos apenas duas palavras (*John* e *father*).

b) Nomes próprios: todos os nomes próprios como, por exemplo, *Paraguay*, *McDonalds*, foram considerados palavras.

c) Palavras produzidas em português: também foram consideradas, por estarmos interessados apenas na quantidade de linguagem produzida. É o caso de “moeda”, “vacinar”.

d) Pausas: elementos que não acrescentam informação relevante ao que está sendo dito (Fortkamp, 1998:217), mas ajudam a situar o desenvolvimento do tópico no contexto geral, particular ou pessoal da conversação (Marcuschi, 2003:62). Foram divididas em:

- pausas lexicais preenchidas ou marcadores lingüísticos verbais lexicalizados (Urbano, 1995). É o caso de: *you know*, *I mean*, *yeab*. Foram consideradas como palavras.
- pausas não-lexicais preenchidas ou marcadores verbais não lexicalizados (Urbano, 1995), como por exemplo: *uh huh*, *mm*, *er*, não foram considerados, pois não carregam informação semântica (Fortkamp, 1998:217).

e) palavras parcialmente produzidas, como “*goi-*”, “*wea*”, “*lo-*”, não foram consideradas.

Vejamos alguns exemplos:

Sb (linha 27): “*this person because when I, when I was a children,*” \_ 10 palavras

Sb (linha 52): “*ah, when I, I a little bit stress, stressed and I*” \_ 10 palavras

Para o cálculo da média, somamos todas as palavras e dividimos pelo tempo total de fala de cada sujeito. Esse mesmo procedimento foi feito quando da realização do teste individual e em duplas.

Em um primeiro momento, por estarmos interessados na quantidade de linguagem produzida, foram consideradas todas as palavras, de acordo



com as especificações anteriores. Logo após, foi observada a qualidade do discurso. Assim, no mesmo período de tempo, verificamos a diversidade do vocabulário, conforme as considerações a seguir.

### 3.2.2. *Diversidade de vocabulário*

A diversidade de vocabulário foi verificada em termos de lexemas diferentes. Para a contagem do número de lexemas diferentes, consideramos apenas os verbos, adjetivos, substantivos e advérbios, conforme as seguintes observações.

a) Verbos: para a contagem dos verbos, consideramos seu aspecto lexical (Huddleston e Pullum, 2002). Vejamos um exemplo: *think* é o lexema e as outras palavras como *thinking*, *thought* são formas ou realizações do lexema, logo, foram contadas apenas como uma ocorrência.

- verbos auxiliares e modais: os verbos auxiliares, como *be*, *have* e os modais (como *will*, *may*, *must*, *shall*, *can*) não foram considerados, com uma ressalva para os verbos “*have*” e “*do*” quando usados como verbos principais da oração, como ocorre nos exemplos:

Sf (linha 106): “*I didn’t have any support for the career*”

Sb (linha 79): “*If you love what you do*”

- os verbos compostos ou complexos, como *look for*, *take off*, *find out*, *go out*, *go through*, no uso idiomático, foram contados como sendo um lexema.

b) Em casos de expressões como *in terms of*, *in fact*, foram considerados apenas os lexemas “*terms*”, “*fact*”. No caso da preposição complexa *in spite of*, apenas “*spite*”.

c) Outras expressões formadas por um verbo + substantivo como *take care*, *pay attention*, foram consideradas como dois lexemas.

d) Advérbios:

- além dos advérbios terminados em *ly*, consideramos também os que modificam os adjetivos e os próprios advérbios.

Não consideramos:

a) Palavras produzidas em português: “moeda”, “vacinar” ou o marcador “né”.

b) Nomes próprios em português (Curitiba, Balneário Camboriú) e em inglês (*Paraguay, Europe, McDonalds, Canada*).

c) Marcadores conversacionais: *you know, well, okay, so, then*.

d) Pronomes, artigos, conjunções, artigos, numerais, preposições (salvo as preposições complexas, conforme já mencionado anteriormente).

Vejamos alguns exemplos:

Sa (linha 73): “*If you think that I’ve been thinking, you know*”. \_ 1 lexema diferente

Sj (linha 42): “*I think ub being a nurse is something difficult, because you need to be very responsible*”. \_ 6 lexemas diferentes.

O cálculo para a média de lexemas diferentes foi basicamente o mesmo do item anterior, ou seja, contamos o número total de lexemas diferentes e dividimos pelo tempo total de fala de cada candidato com a finalidade de obtermos uma média por minuto. O mesmo procedimento foi aplicado para o teste em duplas e o individual.

O outro item a ser observado diz respeito à complexidade gramatical. Para tanto, inicialmente, verificaremos as orações e como essas podem ser.

### 3.2.3. Orações

Há dois tipos básicos de relacionamentos estruturais entre as orações: coordenação e subordinação. Segundo Huddleston e Pullum (2002:1275), a coordenação contrasta com a subordinação. Na subordinação um elemento é o principal e o(s) outro(s) é/são dependente(s). Na coordenação, os elementos apresentam uma relação de igualdade, e normalmente as orações são introduzidas pelas conjunções *and, but, or*.

Ex: *I left the room and Pat followed me*. (coordenada)

*Although the sun is shining, it’s raining in my heart*. (subordinada)

No estudo feito por O’Sullivan (2002:281), a complexidade gramatical foi verificada pelo número total de orações de qualquer tipo divididas

pelo número de sentenças. Porém, devido à dificuldade apresentada em separar as sentenças na fala dos sujeitos, nesta pesquisa, a complexidade gramatical foi considerada em termos de orações subordinadas. Para tanto, verificamos quais as orações subordinadas que ocorreram nas duas situações (candidato-examinador e candidato-candidato-examinador) e contamos suas ocorrências. Foram consideradas todas as orações subordinadas completas, ou seja, não consideramos as orações interrompidas, como por exemplo:

Sd (linha 64): *everything that you-* (oração incompleta – não considerada)

Novamente, para calcularmos a média por minuto, somamos todas as ocorrências de orações subordinadas produzidas no discurso de cada candidato e dividimos pelo tempo total de sua fala. Com a obtenção desses dados, foi-nos possível a construção de tabelas e gráficos.

## 4. Análise dos dados

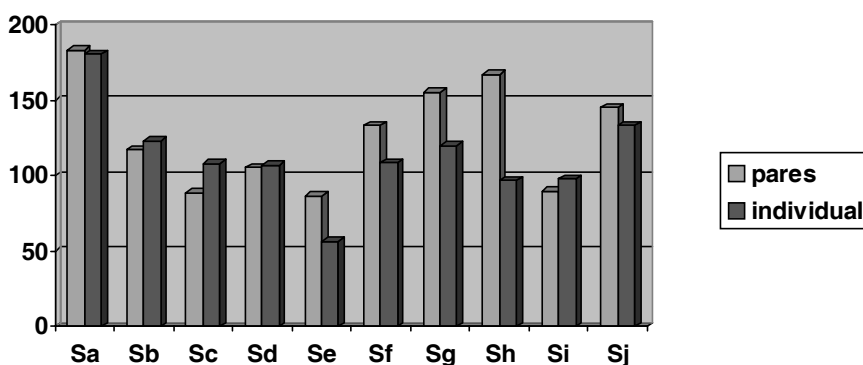
### 4.1. A velocidade da fala

De acordo com as especificações anteriores, foi calculado o número médio de palavras produzidas por minuto pelos sujeitos nas duas diferentes situações (em pares e individual). Foi verificado, com a ajuda de um cronômetro, o tempo (em minutos e segundos) utilizado pelos candidatos nos dois testes e a média final (média da média). Foram ainda calculadas as diferenças existentes nas duas situações por meio da subtração do número médio de palavras em pares (representado pela letra **p**) e individual (representado pela letra **i**). Essa diferença foi também representada em porcentagem. O sinal negativo indica apenas um maior número de palavras produzidas pelo sujeito no teste individual, conforme pode ser observado na Tabela 1. O número médio de palavras por minuto em pares e individual é também representado no Gráfico 1, logo após a Tabela 1.

Tabela 1: Velocidade da fala

Sujeito	Número médio de palavras por minuto em pares (p)	Tempo total pares	Número médio de palavras por minuto individual (i)	Tempo total individual	Diferença (p-i)	Diferença (p-i) em %
Sa	183,33	1:59	180,53	4:31	2,80	1,55
Sb	117,24	2:19	123,20	3:37	-5,96	-4,84
Sc	88,84	2:09	107,80	4:52	-18,96	-17,59
Sd	105,50	2:00	107,20	3:45	-1,70	-1,59
Se	86,70	2:11	56,68	3:22	30,02	52,96
Sf	133,33	1:10	108,49	3:39	24,84	22,90
Sg	155,43	1:45	119,53	3:26	35,90	30,03
Sh	167,07	1:40	96,85	3:10	70,22	72,50
Si	89,53	1:43	97,97	2:57	-8,44	-8,61
Sj	145,22	1:09	133,23	3:06	11,99	9,00
média	127,22	1:63	113,15	3:64	14,07	12,43

Gráfico 1: Velocidade da fala



O Gráfico 1 mostra o número médio de palavras produzidas pelos sujeitos nas duas situações: em pares e individual. O eixo vertical varia de 0 a 200 e representa o número médio de palavras produzidas por minuto pelos sujeitos durante os testes. Pode-se observar que, dos dez sujeitos participantes, quatro (Sb, Sc, Sd, Si) produziram um discurso com um maior número de palavras quando fizeram o teste individualmente. Se verificarmos as diferenças nas duas situações, ou seja, subtraindo-se o número médio de palavras por minuto em pares do número médio de palavras por minuto do teste individual, observamos que Sb produziu 4,84% a mais quando fez o teste individual, Sd, 1,59% e Si, 8,61%. O outro sujeito, Sc, teve uma variação consideravelmente maior, 17,59% a mais por minuto.

Ao analisarmos os demais sujeitos, poderíamos colocá-los em uma ordem crescente de variação, ou seja, da menor para a maior diferença existente entre as duas situações (entre as barras representadas pelo gráfico). Assim, teríamos: **Sa** (1,55 %); **Sj** (9%), **Sf** (22,90%), **Sg** (30,03%), **Se** (52,96%) e **Sh** (72,50%), a favor do primeiro teste.

Além da quantidade de discurso produzido pelos sujeitos, foi verificada também a diversidade do vocabulário, em termos de lexemas diferentes e a complexidade gramatical, em termos de orações subordinadas.

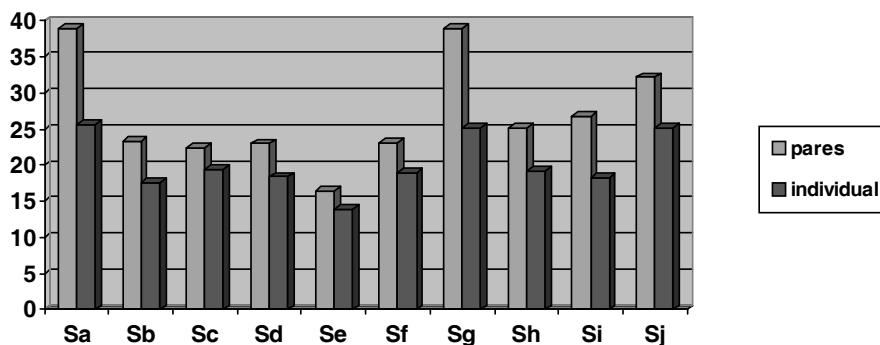
## 4.2. A diversidade do vocabulário

Nesta etapa, o mesmo método foi aplicado, ou seja, contamos os lexemas diferentes, conforme o item 3.2.2, e dividimos o número total pelo tempo falado pelo candidato. A Tabela 2 a seguir mostra ainda o tempo total que cada um dos sujeitos utilizou nas duas situações, a média final do número de lexemas diferentes nas duas situações (média da média) e as diferenças (também em porcentagem) obtidas pela subtração do número médio de lexemas por minuto em pares (representado pela letra **p**) e individual (representado pela letra **i**). O número médio de lexemas diferentes nas duas situações (pares e individual) é também representado no Gráfico 2.

Tabela 2: Diversidade do vocabulário

Sujeitos	Número médio de lexemas diferentes em pares (p)	Tempo total pares	Número médio de lexemas diferentes Individualmente (i)	Tempo total individual	Diferença (p-i)	Diferença %
<b>As</b>	38,89	1:59	25,66	4:31	13,23	51,56
<b>Sb</b>	23,28	2:19	17,68	3:37	5,60	31,67
<b>Sc</b>	22,33	2:09	19,50	4:52	2,83	14,51
<b>Sd</b>	23,00	2:00	18,40	3:45	4,60	25,00
<b>Se</b>	16,51	2:11	13,95	3:22	2,56	18,35
<b>Sf</b>	23,08	1:10	18,90	3:39	4,18	22,12
<b>Sg</b>	38,86	1:45	25,07	3:26	13,79	55,00
<b>Sh</b>	25,15	1:40	19,24	3:10	5,91	30,72
<b>Si</b>	26,74	1:43	18,31	2:57	8,43	46,04
<b>Sj</b>	32,17	1:09	25,16	3:06	7,01	27,86
<b>média</b>	27,00	1:63	20,19	3:64	6,81	33,73

Gráfico 2: Diversidade do vocabulário



Nesse caso, o eixo varia de 0 a 40 e representa o número médio de lexemas diferentes produzidos pelos sujeitos nos dois testes (individual e em pares). O Gráfico 2 demonstrou que todos os sujeitos produziram um discurso com maior número de lexemas diferentes quando interagiram com outro candidato. Ao fazermos a mesma operação anterior (a subtração do número médio de lexemas diferentes em pares pelo número médio de lexemas diferentes individualmente), obtemos resultados que variam de 14,51% a 55%. Se colocássemos em uma ordem crescente, teríamos: Sc (14,51%), Se (18,35%), Sf (22,12%), Sd (25%), Sj (27,86%), Sh (30,72%), Sb (31,67%), Si (46,04%), Sa (51,56%) e Sg (55%), a favor do teste em pares. Podemos observar que a diferença entre as barras é bem expressiva, como é o caso dos sujeitos Si, Sa e Sg. O mesmo aconteceu com relação ao próximo item considerado, a complexidade gramatical.

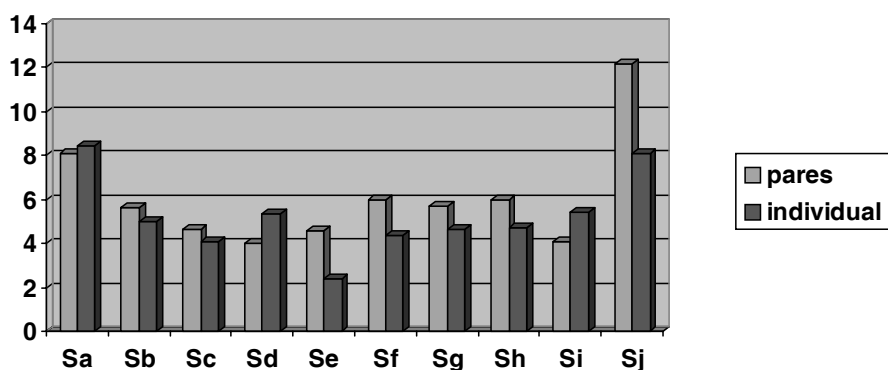
### 4.3. Complexidade gramatical

A complexidade gramatical foi verificada em termos de orações subordinadas, conforme as considerações do item 3.2.3. Fazendo uso do mesmo método utilizado nos dois itens anteriores, obtivemos as seguintes médias e diferenças, representadas na Tabela 3. Novamente, o sinal negativo indica um maior número de orações subordinadas produzidas pelo sujeito no teste individual, conforme pode ser observado na Tabela 3 e no Gráfico 3 a seguir:

Tabela 3: Complexidade gramatical

Sujeitos	Número médio de orações subordinadas em pares (p)	Tempo total pares	Número médio de orações subordinadas Individual (i)	Tempo total individual	Diferença (p-i)	Diferença %
Sa	8,08	1:59	8,41	4:31	0,33	-3,92
Sb	5,60	2:19	4,97	3:37	0,63	12,68
Sc	4,65	2:09	4,11	4:52	0,54	13,14
Sd	4,00	2:00	5,33	3:45	-1,33	-24,95
Se	4,59	2:11	2,37	3:22	2,21	93,67
Sf	5,98	1:10	4,38	3:39	1,60	36,53
Sg	5,71	1:45	4,66	3:26	1,05	22,53
Sh	5,99	1:40	4,73	3:10	1,26	26,64
Si	4,07	1:43	5,42	2:57	-1,35	-24,91
Sj	12,17	1:09	8,06	3:06	4,11	50,99
média	6,08	1:63	5,24	3:64	0,84	16,03

Gráfico 3: Complexidade gramatical



No Gráfico 3, o eixo vertical varia de 0 a 14 e representa o número de orações subordinadas produzidas pelos sujeitos. Pode-se observar que três sujeitos (Sa, Sd, Si) produziram um discurso mais complexo, ou seja, com um número maior de orações subordinadas, quando estavam sozinhos com a examinadora. Novamente, verificam-se as diferenças apresentadas, por meio da subtração do número médio de orações subordinadas produzidas nas duas diferentes situações. Para Sa, a diferença é de apenas 3,92%, ao passo que para Sd, é de 24,95% e para Si 24,91%. Para os demais, se considerarmos a ordem crescente, temos: Sb (12,68%), Sc (13,14%), Sg (22,53%), Sh (26,64%), Sf (36,53%), Sj (50,99%) e Se (93,67%), o que demonstra uma maior complexidade gramatical no primeiro teste realizado, ou seja, no teste em pares.

Se compararmos os três itens - velocidade da fala, diversidade de vocabulário e complexidade gramatical – podemos formar mais uma tabela, mostrando as diferenças médias (em termos do número de palavras, lexe-mas diferentes e orações subordinadas) obtida entre os dois tipos de teste. Entende-se por diferença média o resultado obtido da subtração dos números médios em pares e individuais mostrados nas Tabelas 1, 2 e 3, já comentados anteriormente. Essa diferença foi representada também em porcentagem para uma melhor visualização. O sinal negativo indica um melhor desempenho do sujeito no teste individual. A somatória das diferenças ( $\Sigma di$ ) também aparece na Tabela 4.

Tabela 4: Comparação das diferenças

sujeitos	Velocidade da fala		Diversidade do vocabulário		Complexidade gramatical	
	Diferença	Diferença em %	Diferença	Diferença em %	Diferença	Diferença em %
<b>Sa</b>	2,80	1,55	13,23	51,56	-0,33	-3,92
<b>Sb</b>	-5,96	-4,84	5,60	31,67	0,63	12,68
<b>Sc</b>	-18,96	-17,59	2,83	14,51	0,54	13,14
<b>Sd</b>	-1,70	-1,59	4,60	25,00	-1,33	-24,95
<b>Se</b>	30,02	52,96	2,56	18,35	2,21	93,67
<b>Sf</b>	24,84	22,90	4,18	22,12	1,60	36,53
<b>Sg</b>	35,90	30,03	13,79	55,00	1,05	22,53
<b>Sh</b>	70,22	72,50	5,91	30,72	1,26	26,64
<b>Si</b>	-8,44	-8,61	8,43	46,04	-1,35	-24,91
<b>Sj</b>	11,99	9,00	7,01	27,86	4,11	50,99
$\Sigma di$	140,71	156,31	68,14	322,83	8,39	202,40

Para confirmar se essas diferenças são significativas, usamos um teste estatístico de significação. Nesse caso, foi utilizado o **Teste Não-Paramétrico de Aleatorização** (Guimarães, 2001; Campos, 1983; Siegel, 1956), para se obter a probabilidade exata, sob  $H_0$ , da ocorrência de um resultado qualquer relacionado a pares de dados. Não se exige a verificação da suposição de normalidade. O teste usa basicamente todas as combinações possíveis de resultados ( $2^n$ ), levando em conta o sinal, e verifica quantas somas dessas combinações se encontram acima de um certo valor crítico, valor esse definido pelo nível de significância. Se o resultado observado estiver dentro da região de rejeição, rejeitamos  $H_0$ .

Segundo Brown (1988: 116), o nível de significância é tipicamente definido em  $p < .01$  (1/100) ou em  $p < .05$  (5/100). Para esta pesquisa,



determinamos o segundo caso ( $\alpha = 0,05$ ). Tanto o  $\alpha$  quanto o  $p$  representam essencialmente a mesma coisa, o primeiro símbolo  $\alpha$  é utilizado quando a decisão é inicialmente feita pelo pesquisador e o símbolo  $p$  é usado quando os resultados da pesquisa são relatados.

Assim, para verificar se as diferenças obtidas nos resultados dos testes das duas diferentes situações (pares e individuais) são ou não significativas, foram testadas as seguintes hipóteses, definido como nível de significância  $\alpha = 0,05$  (Levin, 1978: 136):

### 1) Velocidade da Fala

$H1_0$  = não existe diferença entre as situações com relação ao número médio de palavras por minuto.

$H1_1$  = existe diferença entre as situações com relação ao número médio de palavras por minuto.

Encontramos um p-valor = 0,021484, isto é, o resultado se encontra na região de rejeição.

### 2) Diversidade do Vocabulário

$H2_0$  = não existe diferença entre as situações com relação ao número médio de lexemas diferentes por minuto.

$H2_1$  = existe diferença entre as situações com relação ao número médio de lexemas diferentes por minuto

Encontramos um p-valor = 0,001953, ou seja, o resultado se encontra na região de rejeição.

### 3) Complexidade Gramatical

$H3_0$  = não existe diferença entre as situações com relação à complexidade gramatical.

$H3_1$  = existe diferença entre as situações com relação à complexidade gramatical.

Encontramos um p-valor = 0,01953, isto é, o resultado também está na região de rejeição.

Com isso, as hipóteses nulas ( $H_0$ ) foram rejeitadas. Pode-se, portanto, concluir que existe diferença significativa entre as situações para um nível de significância em  $\alpha = 0,05$  (5%), que pode ser observado nas três tabelas anteriores, ou na Tabela 4.

Além disso, um outro ponto a ser considerado foi a relação aluno-teste, ou seja, nosso interesse foi saber como o candidato se sentiu em relação ao teste, conforme veremos a seguir.

#### 4.4. Questionário

Solicitamos aos sujeitos o preenchimento de um questionário constituído de 13 perguntas (abertas e fechadas) ao final do teste individual. As primeiras nove destinavam-se à obtenção de dados pessoais dos sujeitos, como, por exemplo, idade, escolaridade, cursos feitos, experiência fora do país. Já a segunda parte, a nosso ver a mais importante para a pesquisa, formada de quatro perguntas, tinha como objetivo o conhecimento da opinião dos sujeitos sobre as duas diferentes situações do teste (candidato-examinador e candidato-candidato-examinador). Nosso interesse foi saber em qual das duas situações o candidato sentiu mais confiança ou maior dificuldade, como também como ele se sentiu em relação ao teste. Tivemos um total de 10 questionários respondidos, ou seja, todos os sujeitos participantes da pesquisa.

Na leitura das respostas, alguns fatos interessantes foram observados, como, por exemplo, quando é questionado qual a parte mais difícil do teste. As respostas obtidas foram ao mesmo tempo curiosas e um tanto similares. Quatro sujeitos apontaram problemas com relação à própria tarefa solicitada: disseram ter problemas com as gravuras apresentadas durante a realização dos testes. Foram utilizados dois temas: viagens e carreira profissional. Para a pergunta: “na sua opinião, qual foi a parte mais difícil do teste?” algumas respostas foram:

*“Criar comunicação na hora em que você se depara com as figuras”.*

*“Falar, formular as respostas sobre o assunto ao qual não estava preparada para falar”.*

*“Falar com fluência”.*

*“Falar sobre algo que não estava esperando ou que não tinha muito para falar”.*

Ao se perguntar “Como se sentiu em relação aos testes?”, verificou-se que cinco pessoas disseram estar motivadas a fazê-los, sendo que uma delas faz parte de um grupo de pesquisa e gostou bastante da experiência. Duas pessoas disseram estar nervosas ou preocupadas pelo fato de estarem sendo gravadas. Uma delas mostrou-se curiosa por nunca ter participado de um teste com objetivo de pesquisa e duas outras pessoas disseram estar tranquilas.

Uma das principais preocupações quando da formulação do questionário era saber em que situação (sozinho com a examinadora ou com outro candidato) o aluno sentiria mais dificuldade e confiança. Assim, quando os dados foram tabulados, foi observado que, dos 10 sujeitos participantes da pesquisa, oito disseram ter mais dificuldade para discutir o assunto com outro candidato, e sete declararam ter mais confiança quando fizeram o segundo teste, ou seja, quando estavam sozinhos (candidato-interlocutor). Essas respostas foram confirmadas por declarações como:

*“A parte mais difícil do teste foi discutir o assunto com outra pessoa, porque acho mais fácil falar só com a pessoa que está entrevistando...”*  
*“Eu não conhecia o meu parceiro muito bem e por causa disso foi um pouco difícil de interagir”*  
*“Você fica em dúvida se presta mais atenção no que você fala ou no que a outra pessoa diz. Perder a concentração é fatal nessas ocasiões”*  
*“No primeiro teste, estava um pouco nervosa, mas, no segundo, estava à vontade e gostei muito de fazê-lo”*.

Apenas duas pessoas disseram preferir fazer o teste em duplas, conforme podemos observar no comentário de uma delas:

*“Estava tranquila com a outra candidata, mas, quando tive que fazer a entrevista sozinha, fiquei mais nervosa por não ter muito o que falar sobre o que foi perguntado e por não ter a outra pessoa para poder complementar”*.

Com base nesses dados, o resultado obtido foi contrário às conclusões da pesquisa de Humphry-Baker, realizada na Universidade de Manchester, relatada por Taylor (2001:17). Humphry-Baker utilizou-se também de questionários com o objetivo de investigar “como os candidatos se sentem em relação ao teste oral”, ou seja, a reação afetiva dos candidatos para com o teste. Segundo Taylor (2001), Humphry-Baker diz que o teste com outro candidato, ou seja, em pares teve maior preferência entre os seus

sujeitos. Foram apresentadas respostas como: *"I like paired tests."*, *"The test gave me a good opportunity to speak"*.

Em um outro estudo, feito por Egyud e Glover (2001), na Hungria, 14 estudantes responderam a um questionário e revelaram preferir o teste realizado em pares. Egyud e Glover (2001:71) encontraram respostas como: *"It was good because the other candidate could help me."*; *"I think it is very good. It eases the tension."*, *"It was good, but it depends on who one has to be examined with."*, *"It was easier than if I had to do it alone with the examiner"*, *"It was a good idea, since we knew each other and it eased the tension."*, *"It was good because we tried to help each other, though I do not know how it would have felt alone."*.

Egyud e Glover (2001:70) defendem o teste em pares, dizendo que, quando o aluno interage com outro aluno, produz um inglês melhor e que as desigualdades presentes entre os dois são ainda menores do que as presentes entre um aluno e um professor, ou no caso de testes formais, entre um candidato e um examinador.

Esses itens, como a qualidade do discurso e a preferência dos alunos, também foram investigados neste trabalho, conforme pode ser verificado em uma análise mais detalhada a seguir.

## 5. Discussão dos resultados

Iniciando a análise com os dados do sujeito **Sa**, no item velocidade da fala, observamos que seu desempenho nas duas situações (em pares e individual) foi praticamente o mesmo. A diferença, obtida pela subtração do número médio de palavras por minuto em pares pelo número médio de palavras por minuto individual, é de 1,55% a favor do primeiro teste realizado (em pares). Algo semelhante ocorre quando analisamos a diversidade do vocabulário. Nesse caso, o desempenho foi 51,56% superior. Na complexidade gramatical, o sujeito saiu-se melhor individualmente, produzindo um número médio de 8,08 orações subordinadas na primeira situação e 8,41 individualmente: uma diferença de 0,33 ou de 3,92%. Ao verificarmos o último item, a relação aluno-teste, o sujeito diz que o primeiro teste foi mais difícil porque não conhecia o parceiro (**Sb**).

Ao analisarmos as respostas do questionário de **Sb**, observamos que **Sb** não teve a mesma impressão que **Sa**. **Sb** sentiu mais confiança quando estava com outro candidato (no caso **Sa**) e não sentiu nenhuma dificuldade

para fazer os testes. Ainda sobre **Sb**, vemos que o número médio de palavras por minuto, no teste em pares, foi de 117,24 palavras e no individual, de 123,2: uma diferença de 5,96 palavras ou de 4,84%. Entretanto, não podemos dizer que esse sujeito saiu-se melhor quando realizou o teste individualmente porque, ao analisarmos os dois itens restantes, vemos um resultado melhor na interação candidato-candidato. Tanto o número de lexemas diferentes quanto a complexidade gramatical foi beneficiada, conforme já comentado anteriormente.

Um comportamento similar ocorreu com o terceiro sujeito, **Sc**. Este produziu 17,59% palavras a mais quando interagiu apenas com a examinadora. Ao preencher o questionário, **Sc** disse que sentiu mais dificuldade quando interagiu com outro candidato (no caso **Sd**). Ao verificarmos a resposta de **Sd**, notamos que **Sd** também preferiu fazer o teste sozinho com a examinadora. Porém, ao avaliarmos os outros dados (diversidade do vocabulário e complexidade gramatical) de **Sc**, observa-se um melhor desempenho na interação candidato-candidato.

Quanto aos sujeitos **Sd** e **Si**, notamos um fato curioso. Ambos se saíram melhor individualmente em dois itens: velocidade da fala e complexidade gramatical. No caso de **Sd**, a diferença do número médio de palavras produzidas por minuto foi apenas de 1,59%, o que demonstra uma variação muito pequena. **Si** apresenta uma diferença maior, uma média de 8,61% palavras a mais produzidas por minuto quando estava somente com a examinadora. Na complexidade gramatical, observamos quase a mesma diferença, com uma vantagem de 24,95% para **Sd** e 24,91% para **Si**. Na diversidade do vocabulário, ambos produziram um maior número de lexemas diferentes quando interagiram com outro candidato. Na relação aluno-teste, **Si** disse ter tido mais confiança e menor dificuldade quando fez o primeiro teste (candidato-candidato), ao contrário de seu parceiro **Sj** que preferiu fazer o segundo teste (candidato-interlocutor).

Já no caso de **Se**, vemos um melhor desempenho quando realizou o teste em pares. O sujeito produziu um discurso com um maior número de palavras na primeira situação, com uma diferença de 52,96% em relação ao segundo teste. O mesmo ocorre com a complexidade gramatical, uma diferença de 93,67% e de 18,35% com a diversidade de vocabulário. **Se** também diz ter sentido maior confiança e menos dificuldade quando estava com outro candidato, ao contrário de **Sf**, seu parceiro no primeiro teste, o qual disse ter maior confiança e menor dificuldade ao fazer o segundo teste.

Nos sujeitos **Sf** e **Sg**, também observamos um melhor desempenho no primeiro teste realizado. Notamos uma diferença expressiva da velocidade de fala para **Sg**, que pronunciou uma média de palavras de 30,03% a mais por minuto. No outro item, diversidade do vocabulário, **Sg** apresenta uma diferença de 55% lexemas diferentes e **Sf** 22,12%. Quanto à complexidade gramatical, observamos uma diferença de 36,53% para **Sf** e 22,53% para **Sg** quando fizeram o teste em duplas. Pelo relato de **Sg** foi mais difícil fazer o teste com outro candidato. A mesma resposta foi encontrada no questionário de **Sh** – com quem **Sg** fez o primeiro teste.

Quanto ao sujeito **Sh**, notamos um número bem expressivo apresentado no item velocidade da fala. **Sh** produziu um discurso muito mais rápido quando interagiu com outro candidato, com um número médio de palavras de 167,07 e apenas 96,85 individual, uma diferença de 72,5%. Analisando os outros dois itens, a variação do vocabulário e a complexidade gramatical, também pudemos observar um melhor desempenho na interação candidato-candidato.

**Sj** também apresentou um melhor desempenho no primeiro teste realizado. Sua velocidade de fala foi maior (9%), seu discurso demonstrou uma complexidade maior (com uma diferença de 50,99%) e empregou um maior número de lexemas diferentes (27,86%). **Sj** disse que sentiu mais confiança e menos dificuldade quando estava somente com a examinadora, conforme já comentado anteriormente.

Pode-se dizer que, de uma maneira geral, os sujeitos tiveram um melhor desempenho na primeira situação (candidato-candidato-examinador). Com base nesses resultados, foram feitas algumas inferências. O primeiro item (velocidade da fala) mostrou que a maior parte dos sujeitos produziu um discurso com maior número de palavras na situação candidato-candidato-interlocutor. Pode-se concluir que as pausas ou hesitações foram mais curtas e, no caso de uma pausa mais longa, houve a interferência por parte do outro candidato. Pode haver um maior planejamento por parte dos candidatos, ou seja, enquanto um fala o outro “ganha” tempo para planejar o que vai falar. Esse processo evita as repetições ou hesitações e faz com que o discurso soe mais fluente.

Na situação candidato-examinador, apenas quatro candidatos se saíram melhor. Esse comportamento poderia ser explicado devido ao relacionamento com a própria examinadora. Observou-se, em alguns casos, um

discurso mais pausado, mais monitorado por parte dos candidatos nessa segunda situação. Uma atenção maior é dispensada à forma e não tanto à interação na conversa, pois há uma certa hierarquia nessa situação, o que faz com que o examinador seja encarado como um superior. Pode-se dizer que pelo fato de estar sozinho com a examinadora, o candidato pode ficar mais atento à forma. Porém, um discurso com maior número de palavras não quer dizer um discurso bem elaborado. Foram observados, em alguns casos, o uso de palavras em português, repetições, auto-correções. Assim, outros dois aspectos foram também verificados: a diversidade de vocabulário e a complexidade gramatical.

Ao analisarmos o item diversidade de vocabulário, vimos que os sujeitos empregaram um maior número de lexemas diferentes na primeira situação. Algo semelhante ocorreu ao analisarmos a complexidade gramatical. Novamente, os sujeitos produziram um discurso com um maior número de orações subordinadas no primeiro teste realizado.

Entretanto, um fato relevante apresentado foi com relação às respostas fornecidas pelos próprios sujeitos quando do preenchimento dos questionários. Oito dos 10 sujeitos disseram ter mais dificuldade para discutir o assunto com outro candidato e sete declararam ter tido mais confiança quando fizeram o segundo teste, ou seja, quando estavam sozinhos com a examinadora. Esse fato fez com que fizéssemos novas inferências. A primeira seria referente a como foram aplicados os testes: inicialmente os sujeitos fizeram o teste em pares e depois individualmente. Isso pode ter interferido, pois, no segundo teste, talvez os alunos estivessem mais acostumados com os procedimentos, o que fez com que se sentissem mais seguros. Uma outra explicação poderia estar na diferença apresentada pelos candidatos quanto ao nível de proficiência. Um candidato com uma proficiência maior pode preferir fazer o teste sozinho, ao passo que um candidato com o nível de proficiência menor pode preferir fazer o teste com outro candidato.

## **6. Conclusão e encaminhamentos**

O objetivo deste trabalho foi verificar se há diferença no desempenho caso os candidatos façam o teste oral sozinhos ou em pares. Com base nos resultados obtidos, podemos dizer que em ambas as situações (candidato-

candidato-interlocutor e candidato-interlocutor) obtivemos uma amostra lingüística e discursiva dos sujeitos, por meio de tarefas previamente especificadas. Na realidade, são duas situações e dois resultados diferentes. Se levarmos esses resultados para uma sala de aula, por exemplo, teremos que verificar as suas implicações.

Para dizermos qual a melhor situação de teste (em duplas ou individual), antes de mais nada teríamos que verificar as condições da instituição e do ensino. Temos, na maioria dos casos, salas numerosas, poucos recursos, além de outros complicadores como a má-formação de professores. Seria difícil, por exemplo, gravar todos os testes orais de LE e transcrevê-los, tanto pela falta de recursos como pela falta de tempo por parte dos professores. Se levarmos em conta esse fato, poderíamos dizer que seria melhor a realização de um teste oral em duplas, pela maior praticidade. Porém, para essa prática, o professor deveria privilegiar o trabalho em grupo ou em duplas durante as aulas. Afinal, a avaliação e o ensino devem andar juntos. Um outro problema que poderia surgir seria quanto ao score. Avaliar dois alunos ao mesmo tempo é difícil, logo os critérios adotados para a avaliação devem ser muito claros para não afetar a confiabilidade dos resultados.

Em alguns casos, talvez, a confiabilidade do teste poderia ser sacrificada em favor da validade. Em turmas muito numerosas, não é possível o professor fazer um teste oral mesmo em duplas, com duração de, por exemplo, 10 ou 15 minutos e se esse tempo fosse diminuído, poderia afetar a confiabilidade.

Em um instituto de línguas, a situação poderia ser outra. Geralmente as turmas não são tão numerosas, o que poderia fazer com que o professor escolhesse qual a situação que melhor se adaptasse aos seus alunos, mas sempre lembrando que, para isso, o professor teria que ter objetivos bem claros e adaptar sua prática pedagógica. Mas e quanto aos cursos preparatórios para os testes de proficiência? Nesse caso, teríamos que analisar se o efeito retroativo é benéfico ou não. Se o aluno tem como objetivo ser aprovado em um teste de proficiência e consegue se comunicar e utilizar a língua fora daquela situação, ou seja, se há um bom desempenho do aluno, independentemente da preparação para o teste, seria interessante que a oralidade fosse avaliada da mesma maneira que nos testes formais. Assim em um curso preparatório para os exames da UCLES, como o FCE, seria em pares. Já para exames como o IELTS, os alunos seriam avaliados individualmente.



No caso de outras instituições como as universidades, que aplicam testes de proficiência, os professores ou examinadores devem também analisar as variáveis contextuais, tais como o propósito do teste e o uso que pretendem fazer dele e de seus resultados. Além disso, devem ser observadas questões como as condições de ensino, os objetivos da instituição e dos próprios alunos, a prática pedagógica, os resultados de pesquisas científicas ou empíricas e a partir de todos esses dados escolher a melhor situação (pares ou individual) que deve ser aplicada naquela ocasião específica.

Na realidade, dispomos de duas situações com dois resultados diferentes, que indicam ter havido uma tendência por parte dos sujeitos em produzirem um discurso com maior número de palavras, mais lexemas diferentes e uma maior complexidade gramatical quando interagiram com outro candidato. Porém, oito dos dez sujeitos participantes da pesquisa disseram ter mais confiança quando fizeram o teste individualmente. Esse fato nos leva a crer que seria interessante ouvir mais nossos alunos para procurar saber o porquê dessa preferência. Poderíamos tentar explicar isso pelo medo que o aluno tem de se sentir prejudicado caso seu parceiro tenha um nível de proficiência diferente do seu, ou ainda, conforme já atestado por Hughes (2003:121), o desempenho de um candidato pode ser afetado pelo do outro. É o que pode ocorrer, por exemplo, se um candidato dominar a fala e não permitir que o outro mostre o que pode fazer; ou ainda, quando um candidato acha que já falou o suficiente e fica esperando o outro se manifestar.

Um outro fato também poderia ser a falta de conhecimento por parte do professor e do aluno com relação a pesquisas que envolvem as duas situações. De acordo com Hughes (2003:119), na situação individual, o candidato não toma iniciativa, apenas responde as perguntas, é como se o candidato falasse para um superior. Já na situação candidato-candidato-examinador, há uma maior negociação na comunicação entre os participantes, o que mostra variações mais freqüentes na linguagem utilizada. Seria interessante que o professor informasse aos seus alunos que a situação de interação entre os candidatos gera um desempenho melhor na maior parte dos casos. Além disso, segundo Underhill (1987:45), essa situação é a mais natural do mundo – duas pessoas conversando sobre um assunto de interesse comum. Porém, é algo não muito fácil em um teste oral, pois ambos os candidatos devem estar, de certa forma, relaxados e confiantes

para iniciarem e desenvolverem uma conversa. O sucesso depende em grande parte da habilidade do interlocutor (ou, no caso, do professor) em criar um ambiente adequado.

Recebido em julho de 2005

Aprovado em abril de 2007

E-mail:

## REFERÊNCIAS BIBLIOGRÁFICAS

- BROWN, J. D. 1988. *Understanding research in second language learning: a teacher's guide to statistics and research design*. Cambridge: Cambridge University Press.
- CAMPOS, H. 1983. *Estatística experimental não-paramétrica*. Piracicaba: ESALQ/USP.
- EGYUD, G.; GLOVER, P. 2001. Oral testing in pairs – a secondary school perspective. *English Language Teaching Journal*, 55, n. 1: 70-76.
- FOOT, M. 1999. Relaxing in pairs. *English Language Teaching Journal*, 51, n. 1: 36-41.
- FORTKAMP, M. B. M. 1998. Measures of working memory capacity and L2 oral fluency. *Ilha do Desterro*, 35 jul./dez: 201-238. Florianópolis.
- GUIMARÃES, P. R. B. 2001. *Estatística não-paramétrica*. Curitiba. Apostila (Disciplina Estatística Não Paramétrica) – Curso de Estatística, Setor de Ciências Exatas, Universidade Federal do Paraná.
- HUDDLESTON, R.; PULLUM, G. K. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- HUGHES, A. 2003. *Testing for language teachers*. 2. ed. Cambridge: Cambridge University Press.
- LAZARATON, A. 2002. *A qualitative approach to the validation of oral language tests*. Cambridge: Cambridge University Press.
- LENNON, P. 1990. Investigating fluency in EFL: a qualitative approach. *Language Learning*, 40, n. 3: 387-417.
- LEVIN, J. 1978. *Estatística aplicada a ciências humanas*. São Paulo: Harbra.
- MARCUSCHI, L. A. 2003. *Análise da conversação*. 5. ed. São Paulo: Ática.
- O'SULLIVAN, B. 2002. Learner acquaintanceship and oral proficiency test pair-task performance. *Language Testing*, 19, n. 3: 277-295.

- RADFORD, A. 1997. *Syntactic theory and the structure of English: a minimalist approach*. Cambridge: Cambridge University Press.
- ROSS, S.; BERWICK, R. 1992. The discourse of accommodation in oral proficiency examinations. *Studies in Second Language Acquisition*, 14: 159-176.
- SCARAMUCCI, M. V. R. 2000. Proficiência em LE: considerações terminológicas e conceituais. In SIGNORINI, I.; OTTONI, P.; FIAD, R. S. (Org.). *Trabalhos em Lingüística Aplicada*, 36: 11-22, Campinas, jul./dez.
- SIEGEL, S. 1956. *Estatística não-paramétrica para as ciências do comportamento*. Rio de Janeiro: McGraw-Hill.
- TAYLOR, L. 2001. The paired speaking test format: recent studies. *Research Notes*, n. 6:15-17. Cambridge.
- UNDERHILL, N. 1987. *Testing spoken language: a handbook of oral testing techniques*. Cambridge: Cambridge University Press.
- UNIVERSITY OF CAMBRIDGE LOCAL EXAMINATIONS SYNDICATE. 1999. *English for life: an introduction to Cambridge EFL examinations*. Cambridge: UCLES.
- \_\_\_\_\_. 2003. *FCE Handbook: First Certificate in English*. Cambridge: UCLES.
- \_\_\_\_\_. 2003. *Instructions to oral examiners: March 2004-December 2005*. Cambridge: UCLES.
- URBANO, H. 1995. Marcadores conversacionais. In: PRETI, D. *Análise de textos orais*. São Paulo: FELCH/USP: 81-101.
- YOUNG, S; MILANOVIC, M. 1992. Discourse variation in oral proficiency interviews. *Studies in Second Language Acquisition*, 14, n. 4: 403-424.