

PRODUÇÃO ESCRITA EM LÍNGUA PORTUGUESA E CASTELHANA: REALIZAÇÃO DE IMIGRANTES ESTABELECIDOS EM SÃO PAULO (1960 - 1970)

María Del Pilar Sacristán Martín

I- INTRODUÇÃO

Esta comunicação resume a Dissertação de Mestrado apresentada ao Departamento de Linguística e Línguas Orientais da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, e defendida em dezembro de 1985.

Trata-se de uma pesquisa em linguística comparada, na qual se explora a aquisição e domínio da língua portuguesa por imigrantes castelhanos e a manutenção da primeira língua destes falantes. Procurou-se estudar comparativamente o castelhano e o português, através do uso efetivo da linguagem destes imigrantes, e verificar se o tempo de permanência no país constituía variável significativa no desempenho destes sujeitos em ambas as línguas.

A descrição de dois sistemas linguísticos confrontados, como o propõe a Análise Contrastiva (Lado, 1973 e outros), é uma perspectiva válida e muito útil nos estudos comparativos, não obstante, apresenta limitações que só podem ser superadas levando-se em conta a produção real ou fala dos indivíduos. O desempenho adequado dos bilingües não decorre apenas do conhecimento funcional das línguas, mas da utilização apropriada dos três níveis: sistema, norma e uso. Considere-se também que os problemas linguísticos do bilingüe ou do aprendiz não podem ser atribuídos apenas ao fenômeno interferencial.

No ensino de línguas estrangeiras, e de um ponto de vista pragmático, é relevante, para a abordagem comparativa, o conhecimento do comportamento linguístico do aprendiz. Neste sentido, por manipular situações reais, a Análise de Erros (Richards, 1974 e outros) facilita a obtenção de resultados frutíferos e de aplicação relativamente rápida. O estudo cuidadoso da produção verbal do indivíduo possibilita extrair dados efetivos para a determinação de prioridades no ensino de língua estrangeira e a seleção e preparação de material didático adequado, gradativo e objetivo, fundamentado em bases concretas. O erro, por ser entendido como um passo necessário no processo pedagógico, permite conhecer as dificuldades do aprendiz e subsaná-las, bem como reconhecer as diferentes etapas pelas quais passa o indivíduo na aquisição de outra língua.

Com a perspectiva descrita, este trabalho analisa a produção escrita em língua portuguesa e castelhana de dois grupos de imigrantes estabelecidos em São Paulo, na década de 1960 e 1970. Para a autora do presente estudo, urge realizar pesquisas que detectem, descrevam e analisem os erros produzidos por sujeitos que manipulam o português e o castelhano. Tais trabalhos fornecerão informações valiosas para, num futuro, reconhecer os erros específicos, passíveis de ocorrer sistematicamente no desempenho dos indivíduos, e correlacioná-los com os diferentes tipos de sujeitos (os que têm por língua materna o castelhano, os que têm o português por primeira língua e os estudantes de ambas). Os dados obtidos facilitarão aos profissionais da área a elaboração de estratégias para corrigir, aperfeiçoar e dinamizar o ensino e a aprendizagem do castelhano e do português, além de auxiliar na tarefa de formação e de atuação do tradutor, seja ele profissional ou não.

A carência de pesquisas, de bibliografia (estudo contrastivo em português e castelhano), e de referencial teórico adequadamente sustentado por dados, fazem com que este trabalho tenha um caráter pioneiro. Justamente por tratar-se de uma primeira ten-

tativa, ainda que lingüísticamente coerente e operacional de análise e sistematização do corpus, apresenta as limitações daí decorrentes; no entanto, pretende estimular reflexões, abrir perspectivas e incentivar novas pesquisas nesta linha.

II- OBJETIVOS

Os objetivos específicos da pesquisa podem ser formulados como segue.

- a) Verificar a influência da variável tempo de permanência no país na produção escrita em língua portuguesa e em língua castelhana de sujeitos castelhanos estabelecidos em São Paulo na década de 60 e 70.
- b) Analisar esta produção escrita de acordo com as seguintes categorias e sub-categorias:
 - i) Nível Paradigmático e nível Sintagmático:
 - ii) Inadequação: Sem Interferência
e
Com Interferência
 - iii) Erro: Sem Interferência
e
Com Interferência: Empréstimo
e
Hibridismo
- c) Comparar a produção nas duas línguas nas mesmas categorias, em cada grupo e entre grupos.

III- METODO

1. Sujeitos

Reuniu-se uma amostra de informantes com as seguintes características: a) imigrantes cuja única língua materna fosse o castelhano e com instrução escolar básica efetuada na Espanha; b) chegados ao Brasil em idade adulta (por já terem incorporado a língua materna) em duas épocas diferentes mas suficientemente próximas, de forma que ainda mantivessem certa fluência no castelhano; e c) do sexo masculino pelas oportunidades de relacionamento social e profissional mais amplas que as do sexo feminino em geral.

Foram entrevistados 12 sujeitos do sexo masculino, espanhóis, naturais de Castela, estabelecidos no Brasil, em São Paulo, seis deles entre os anos de 1959 e 1962 (Grupo 1), e os seis restantes entre 1969 e 1972 (Grupo 2).

2. Material

Foi apresentada como estímulo aos sujeitos para que redigissem em ambas as línguas uma prancha de Peabody (1968). Trata-se de uma gravura em cartolina plastificada, representando uma cena familiar: uma mulher sentada numa poltrona, dando a mamadeira a um bebê, rodeada por um menino e uma menina de aproximadamente dez anos e um homem.

3. Procedimento de coleta

Para a coleta do corpus, cada informante passou pelas seguintes etapas:

- a) contato telefônico inicial, e outros posteriores quando se fez necessário;
- b) primeira entrevista, controlada pela pesquisadora conforme a seqüência: i) conversa informal de caráter pessoal; ii) perguntas para identificação do informante, através de um questionário especialmente elaborado para este fim; e iii) apresentação da prancha e produção do texto em determinado idioma; e

- c) segunda entrevista, na qual foi novamente apresentada a mesma prancha e solicitado novo texto, agora no outro idioma, observando-se um intervalo de no mínimo 15 dias entre uma entrevista e outra.

Convém ressaltar que sempre foi omitido aos sujeitos o caráter linguístico da pesquisa, e que procedeu-se a um sorteio para evitar o possível viés resultante de uma mesma seqüência de utilização do idioma.

4. Fidedignidade das avaliações

O corpus foi transcrito "ipsis litteris" dos manuscritos e, por não haver ilegibilidade, foi desnecessária a confrontação por parte dos juizes naquele momento.

O corpus em castelhano foi avaliado em Madri, por três juizes, professores universitários dedicados ao ensino e à pesquisa em castelhano. O corpus em português foi igualmente avaliado por três juizes, em São Paulo, professores universitários de língua portuguesa e linguística.

Pediu-se aos juizes que corrigissem os textos como se fossem uma redação de alunos, e que assinalassem simplesmente os elementos que consideravam inadequados ou incorretos podendo fazer as observações que considerassem pertinentes nas margens da folha.

Para estabelecer a fidedignidade fez-se o cômputo de acordos e desacordos com relação aos elementos assinalados pelos juizes e aplicou-se o procedimento de cálculo sugerido por Bomtempo (1975). O índice de concordância foi de 84,03% e 79,5% para o corpus em português e em castelhano respectivamente. Um segundo cálculo de fidedignidade, classificando os desacordos em relevantes (transgressão efetiva não assinalada por algum juiz) e não relevantes (questões de uso, de estilo e de redação "strictu sensu" não assinalada por algum juiz) alcançou uma porcentagem de concordância de 99,03% e 97,91% para o corpus em castelhano e português respectivamente.

5. Procedimento de análise

Por não ter sido encontrado na literatura específica um modelo de classificação que se adaptasse a esta pesquisa, fez-se necessário elaborar uma categorização especial que preenchesse três requisitos básicos: a) refletir com a necessária fidelidade, os diferentes tipos de transgressões extraídos do corpus; b) manter dentro dos limites possíveis, uma coerência linguística na classificação; e c) permitir a emergência de um resultado operacional, isto é, obter no final dados que pudessem ser apreendidos com relativa facilidade pelos interessados no ensino das línguas abordadas. A medida que o material linguístico produzido permitia inferir opções classificatórias, foram sendo selecionadas aquelas que respeitavam os critérios acima estabelecidos.

Optou-se pela lexia (Pottier, 1977) como unidade classificatória básica e ponto de partida para a categorização dos desvios, pois tal conceito revelou-se operacional para manipular o corpus.

Da observação do material coletado inferiu-se que determinados desvios limitavam-se à lexia em si mesma, e que a substituição pela unidade lexical apropriada correspondente corrigia naturalmente o contexto. No entanto, outras infrações ultrapassavam os limites da lexia, incidindo em outras lexias, no sintagma e no enunciado. A substituição nestes casos, ou não era possível (presença x ausência e vice-versa) ou então não se adequava às restrições de substituição que foram elaboradas especialmente para operacionalizar a distinção em pauta. Utilizando a denominação nível Paradigmático, agruparam-se nesta rubrica todos os desvios que permitiram a substituição descrita. A designação nível Sintagmático reservou-se para as transgressões que afetavam o contexto de tal forma que a substituição não era possível, a não ser modificando outros elementos adjacentes.

Após a primeira grande distinção (nível Paradigmático e Sintagmático) o material coletado permitiu uma nova classificação : Inadequação e Erro. Embora não se possam estabelecer limites rígidos neste "continuum", para operacionalizar a dicotomia, optou-se por incluir na categoria Inadequação as emissões desviadas que não incidiam no sistema mas que tinham sido detectadas pelos juízes por: a) sua pouca frequência, b) porque regras prescritivas de caráter predominantemente estilístico as desaconselhavam, e c) por pertencerem a outros registros de linguagem. Considerou-se Erro toda a transcrição que comprometia: a) as regras funcionais do idioma em questão e b) o conjunto de traços particulares e convencionais que caracterizam o comportamento deste idioma.

A cada desvio produzido (Erro ou Inadequação) verificou-se se tal infração estava relacionada com a língua de partida ou não, no primeiro caso, definiu-se o desvio Com Interferência. Quando não se constatava nenhuma relação com a Língua de Partida (LP) (1), ou quando não era possível afirmar com segurança que decorria de interferência da LP foi denominado Sem Interferência.

O Erro Com Interferência mereceu uma segunda identificação. Nos casos em que o elemento havia sido tomado literalmente (forma e conteúdo) da LP foi denominado Empréstimo; se apresentava uma mescla de ambas as línguas (forma e conteúdo) foi designado por Hibridismo.

Todos os desvios foram identificados de acordo com a classificação acima exposta e efetuou-se a seguir, a fidedignidade da categorização, com a participação de um juiz, professor doutor em lingüística. Foi obtido um índice de 100% de fidedignidade, e os 17 desvios onde não houve acordo foram discutidos até que se chegou a um consenso.

O esquema abaixo permite uma visualização clara do quadro classificatório descrito.

| | | | |
|---|----------------|-------|--|
| NÍVEL PARADIGMÁTICO (FORMA E CONTEÚDO) | 1. inadequação | | sem interferência (PI-1) |
| | | | com interferência (PI-2) |
| | 2. erro | | sem interferência (PE-1) |
| | | | com interferência empréstimo (PE-2A) |
| | | | hibridismo (PE-2B) |

| | | | |
|--|----------------|-------|--|
| NÍVEL SINTAGMÁTICO (FORMA E CONTEÚDO) | 1. inadequação | | sem interferência (SI-1) |
| | | | com interferência (SI-2) |
| | 2. erro | | sem interferência (SE-1) |
| | | | com interferência empréstimo (SE-2A) |
| | | | hibridismo (SE-2B) |

(1) Língua de Chegada (LC): língua em que o sujeito produziu o texto.

Língua de Partida (LP): o outro idioma que o sujeito manipula e que pode interferir ou não na sua produção verbal.

IV- RESULTADOS1. Análise qualitativa

Os desvios extraídos do corpus e classificados como Inadequação caracterizaram-se pela inobservância de regras prescritivas de natureza estilística da LC e pela utilização de elementos lingüísticos existentes na LC mas de uso pouco frequente, raro ou arcaico no contexto em que apareceram ou no espectro geral do idioma em questão. Foi detectado também o uso de registros de fala discrepantes do nível de língua em que o indivíduo se expressou. Algumas ocorrências, neste caso, estavam associadas a padrões da LP outras não.

As Inadequações Sem Interferência caracterizaram-se em ambas as línguas e níveis pela escolha ou combinatória formal de signos existentes na LC com traços morfo-sintáticos e/ou semânticos inconvenientes para o ambiente lingüístico em que foram plasmados. A Forma Produzida (FP) pelo sujeito correspondeu a um signo existente na LC porém impróprio para o ambiente lingüístico em que se manifestou. Por exemplo:

- (p) "um vestido cuadriculado" por "um vestido xadrez" (PI-1)
 (c) "...pero que más tarde se recuerda con tanto cariño ..." por
 "...pero que más tarde se recordará con tanto cariño..." (SI-1)

As Inadequações Com Interferência, em ambas as línguas, consistiram na seleção ou combinatória formal de elementos lingüísticos existentes na LC e na LP, porém privilegiando traços morfo-sintáticos e/ou semânticos específicos da LP, em detrimento da LC. Estas ocorrências parecem estar vinculadas às semelhanças ou identidades entre as faces do signo, nos dois idiomas. Os signos marcados(2) da LC, semelhantes ou idênticos aos não marcados da LP foram preferidos em detrimento dos signos não marcados. Este fenômeno ocorreu fundamentalmente quando a LC possuía dois signos correlatos em distribuição complementar ou parcialmente equivalentes, um deles não marcado e o outro marcado; e a LP um signo correspondente não marcado (independente da existência ou não do signo marcado). Por exemplo:

- (p) "... a mim no me traz..." por "...nao me traz..." (SI-2)
 (c) "observamos un cuadro nítidamente familiar..." por "...observamos um cuadro claramente familiar..." (PI-2)

Na categoria Erro, classificaram-se todos os desvios que incidiram no nível funcional da língua e nos modelos convencionais ortográficos, morfo-sintáticos, léxicos e semânticos admitidos pela gramática da LC. No Erro Sem Interferência a LP não demonstrou ter participação direta; sua participação, entretanto, parece existir de forma indireta, pela vacilação do bilingüe na manipulação e discriminação dos modelos da LC - fonte de grande número de desvios incomuns aos monolíngües.

Nesta categoria encontraram-se dois tipos básicos de respostas: a) as que transgrediam regras idênticas nos dois idiomas (cf. concordância) e que sugerem falta de conhecimento sistemático de ambas as línguas; e b) as que transgrediam regras diferentes ou sem equivalência nestas línguas. Estas últimas parecem fomentar um procedimento de subdiferenciação (ausência da manifestação lingüística da regra da LC) ou de superdiferenciação (aplicação da regra em todas as situações). Tanto em português como em castelhano o Erro Sintagmático Sem Interferência foi bastante significativo e incluiu, especialmente, desvios do tipo a) antes citado (omissão de verbos e conjunções; ausência, presença imprópria ou troca de preposições; concordância, etc.).

No nível Paradigmático em castelhano, encontraram-se respostas similares às que se verificam também nos monolíngües (separação indevida de lexias, transcrição ortográfica incorreta, etc.) Porém, no nível Paradigmático em português, detectou-se um

- (2) Denomina-se elemento marcado àquele que apresenta particularidades formais e/ou semânticas especiais, tanto no nível morfo-sintático quanto léxico, e que se distingue ou opõe a outro elemento correlato (não marcado). Assim "seguridade" em português considerar-se-á marcado com relação a "segurança", pois sua presença num ato lingüístico tem conotação especial. Para o castelhano "reminiscencias" é signo marcado com relação a "recuerdos", não marcado.

número significativamente alto de desvios relacionados com o desconhecimento ortográfico-fonológico do português, entre eles: a) problemas de consonantismo, em razão da falta de correspondência existente entre os sistemas fonológico e ortográfico das duas línguas quanto às consoantes fricativas linguoalveolares (Ex.: transcrição de s por z para o fonema /z/: "organizações" por "organizaçoes"); b) problemas de vocalismo, pela discrepância entre os sistemas vocálicos dos dois idiomas e da alofonia vocálica do português do Brasil, especialmente a neutralização entre /o/ e /u/ e entre /e/ e /i/ (ex.: transcrição de ai por ãe para o ditongo nasal /ãy/: "mai" por "mãe", transcrição de u por o para a vogal final reduzida /u/: "continuu" por "continuo"); e c) problemas de grafia fonológica ("mais" por "mas") e falsas analogias ("pães" por "pais").

Os Erros Com Interferência por Empréstimo privilegiaram em ambas as línguas o nível Paradigmático. Nesta categoria incluíram-se todas as ocorrências que apresentaram forma literal na LP. Os resultados sugerem que a semelhança formal entre os signos da LP e da LC facilita a transferência do equivalente da LP, mas a semelhança formal não é condição "sine qua non" para tal transposição. Por exemplo:

- (p) "persona" por "pessoa" (PE-2A)
- (c) "mamadera" por "biberón" (PE-2A)

Ratifica-se no corpus a noção de que é relevante a frequência de uso de certos signos na experiência lingüística quotidiana dos sujeitos. Também é possível que a lei da economia lingüística esteja presente nestes casos.

A face significativa do signo parece ter demonstrado mais força do que a significante. No caso de unidades formalmente diferentes, revelou-se o procedimento de transferência, no caso de unidades formalmente semelhantes pode estar implicada a subdiferenciação semântica do signo da LC. É importante observar que no corpus em português constatou-se uma quantidade significativa de empréstimos de lexias gramaticais (verbo ser, preposições, artigos, pronomes pessoais, etc.), tema que será tratado na conclusão deste trabalho.

A ocorrência de Erros Com Interferência por Híbridismo foi relativamente baixa, comparada com os outros tipos de Erros, porém apresentou uma grande variedade de mesclas de forma e conteúdo. Nas emissões deste gênero verificaram-se desde a transcrição da forma da LC com apenas um traço correspondente ao padrão ortográfico da LP, no início, no meio ou no final do signo, até a transcrição do morfema lexical da LP acrescido do morfema gramatical de grau diminutivo da LP, transcrito na grafia da LC. Por exemplo:

- (p) "eslabão" por "elo" (PE-2B)
- (p) "casalsilho" por "casalzinho" (PE-2B)
- (c) "en cuanto" por "mientras" (PE-2B)
- (c) "...mismo con su adolescente edad..." por "...a pesar de..." (SE-2B)

2. Análise quantitativa

Para o tratamento estatístico, num primeiro passo, efetuou-se o registro das ocorrências discriminadas por categorias e por idioma, com o cálculo das porcentagens. Apresentam-se aqui as tabelas 1, 2, 3 e 4 e as figuras 1 e 2, onde aparece o desempenho dos dois grupos, separado por categorias, níveis e idiomas.

A segunda etapa desta análise está constituída pelos testes estatísticos aplicados ao corpus. Considerando-se as limitações físicas desta comunicação e com o intuito de expor com maior detalhe os resultados obtidos, omitir-se-ão as tabelas e figuras correspondentes aos citados testes, bem como os procedimentos técnicos propriamente estatísticos e descritivos. Porém, estes dados encontram-se no texto original, à disposição dos que os desejarem consultar.

TABELA 1

Desempenho de G1 e G2 no corpus em língua portuguesa:
Nível Paradigmático

| GRUPOS CATEGORIAS | | G 1 | | G 2 | | TOTAL | |
|---|-------|-----|------|-----|------|-------|------|
| | | F | % | F | % | F | % |
| I N A D E Q U A Ç A O | PI-1 | 1 | 1,2 | 4 | 3,2 | 5 | 2,3 |
| | PI-2 | 1 | 1,2 | 6 | 4,8 | 7 | 3,3 |
| E R R O | PE-1 | 32 | 37,2 | 25 | 19,8 | 57 | 26,9 |
| | PE-2A | 43 | 50,0 | 71 | 56,3 | 114 | 53,8 |
| | PE-2B | 9 | 10,4 | 20 | 15,9 | 29 | 13,7 |
| TOTAL | | 86 | 100 | 126 | 100 | 212 | 100 |

TABELA 2

Desempenho de G1 e G2 no corpus em língua portuguesa:
Nível Sintagmático

| GRUPOS CATEGORIAS | | G 1 | | G 2 | | TOTAL | |
|---|-------|-----|------|-----|------|-------|------|
| | | F | % | F | % | F | % |
| I N A D E Q U A Ç A O | SI-1 | 0 | 0,0 | 1 | 2,8 | 1 | 1,5 |
| | SI-2 | 3 | 9,4 | 3 | 8,6 | 6 | 8,9 |
| E R R O | SE-1 | 19 | 59,4 | 23 | 65,7 | 42 | 62,7 |
| | SE-2A | 6 | 18,7 | 5 | 14,3 | 11 | 16,4 |
| | SE-2B | 4 | 12,5 | 3 | 8,6 | 7 | 10,5 |
| TOTAL | | 32 | 100 | 35 | 100 | 67 | 100 |

TABELA 3

Desempenho de G1 e G2 no corpus em língua castelhana:
Nível Paradigmático

.8.

| GRUPOS CATEGORIAS | | G 1 | | G 2 | | TOTAL | |
|---|-------|-----|------|-----|------|-------|------|
| | | F | % | F | % | F | % |
| I N A D E Q U A Ç A O | PI-1 | 2 | 3,7 | 2 | 5,6 | 4 | 4,5 |
| | PI-2 | 6 | 11,1 | 11 | 30,6 | 17 | 18,9 |
| E R R O | PE-1 | 11 | 20,4 | 9 | 25,0 | 20 | 22,2 |
| | PE-2A | 32 | 59,3 | 7 | 19,4 | 39 | 43,3 |
| | PE-2B | 3 | 5,5 | 7 | 19,4 | 10 | 11,1 |
| TOTAL | | 54 | 100 | 36 | 100 | 90 | 100 |

TABELA 4

Desempenho de G1 e G2 no corpus em língua castelhana:
Nível Sintagmático

| GRUPOS CATEGORIAS | | G 1 | | G 2 | | TOTAL | |
|---|-------|-----|------|-----|------|-------|------|
| | | F | % | F | % | F | % |
| I N A D E Q U A Ç A O | SI-1 | 4 | 12,1 | 3 | 10,4 | 7 | 11,3 |
| | SI-2 | 8 | 24,2 | 5 | 17,2 | 13 | 20,9 |
| E R R O | SE-1 | 16 | 48,5 | 13 | 44,8 | 29 | 46,8 |
| | SE-2A | 4 | 12,1 | 3 | 10,4 | 7 | 11,3 |
| | SE-2B | 1 | 3,1 | 5 | 17,2 | 6 | 9,7 |
| TOTAL | | 33 | 100 | 29 | 100 | 62 | 100 |

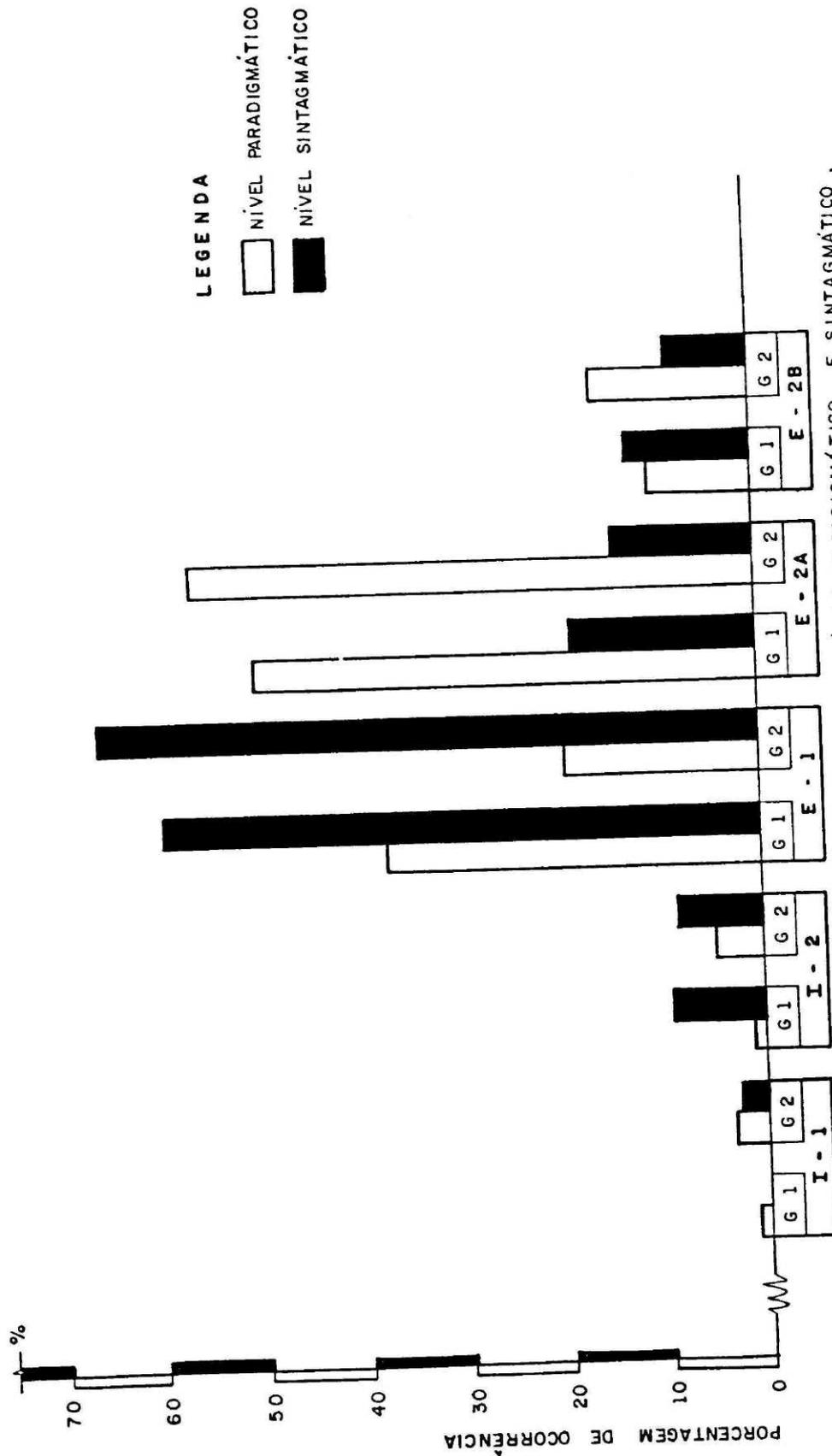


FIG. 1 - PORCENTAGEM DE OCORRÊNCIA DAS CATEGORIAS NOS NÍVEIS PARADIGMÁTICO E SINTAGMÁTICO, DO CORPUS EM LÍNGUA PORTUGUESA, PARA G1 E G2

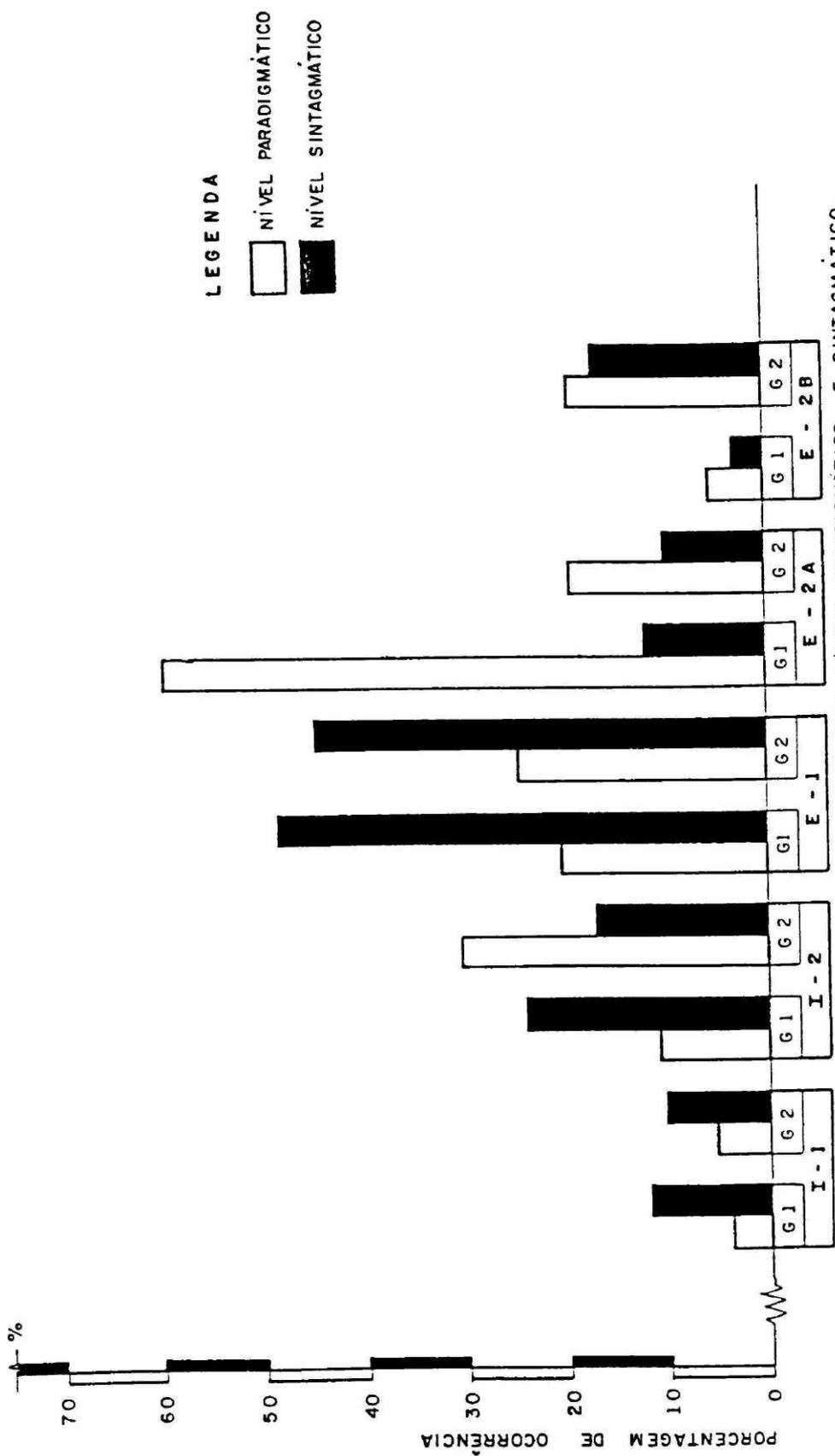


FIG. 2 - PORCENTUAL DE OCORRÊNCIA DAS CATEGORIAS NOS NÍVEIS PARADIGMÁTICO E SINTAGMÁTICO, DO CORPUS EM LÍNGUA CASTELHANA, PARA G1 E G2

a) Resultado do cálculo das porcentagens

A seguir expõe-se de forma sucinta as constatações básicas extraídas do resultado do cálculo de porcentagens, nas duas sub-categorias gerais.

i) Inadequação

Os dados revelaram que G1 e G2 produziram maior número de Inadequações no idioma castelhano do que no português; e em castelhano, o nível Sintagmático superou o Paradigmático. No corpus em português ambos os grupos apresentaram níveis muito baixos de Inadequação, e o nível Paradigmático foi ligeiramente inferior ao Sintagmático. Tanto em português quanto em castelhano a Inadequação Sem Interferência obteve índices inferiores comparada à Inadequação Com Interferência. Em língua portuguesa a Inadequação Com Interferência, no nível Sintagmático obteve maiores ocorrências em ambos os grupos do que no nível Paradigmático. Em língua castelhana, a afirmação anterior foi verdadeira para G1, porém, para G2 o número mais elevado de ocorrências constatou-se na categoria PI-2, que obteve 30,6%, a maior porcentagem da categoria Inadequação.

ii) Erro

A categoria Erro concentrou o maior número de respostas nos dois idiomas, comparativamente com a categoria Inadequação. O Erro Sem Interferência obteve índices altos e relativamente aproximados em ambas as línguas e grupos, e o nível Sintagmático superou, nas duas línguas o nível Paradigmático.

Os Empréstimos tiveram maior incidência no nível Paradigmático do que no Sintagmático, em ambas as línguas, e foram preferidos ao Hibridismo. O Erro Com Interferência por Hibridismo, em português, nos dois níveis, apresentou porcentagens equivalentes e baixas em ambos os grupos. Não obstante, em língua castelhana, o Hibridismo foi mais significativo em G2 do que em G1.

Globalmente o comportamento de G1 e G2 não difere quanto à concentração das ocorrências na categoria Erro. No corpus em português, em ambos os grupos, a categoria Erro Sintagmático Sem Interferência forneceu a maior concentração de respostas. G1 apresentou mais Erros Paradigmáticos Sem Interferência em língua portuguesa do que em castelhano. Ao contrário, G2 ofereceu mais respostas nesta sub-categoria em castelhano do que em português.

Enquanto G1 registrou um número elevado de Empréstimos Paradigmáticos em ambas as línguas, e superior em castelhano, G2 efetuou mais Empréstimos ao redigir em português. O Hibridismo obteve valores semelhantes em castelhano e em português para G2. O grupo 1 ofereceu níveis muito baixos nesta sub-categoria em língua castelhana, e um pouco mais elevados no português.

b) Testes estatísticos

Numa segunda etapa da análise quantitativa, efetuaram-se os cinco testes estatísticos que a seguir se resumem.

Aplicou-se primeiramente um χ^2 que procurou verificar se a ocorrência das sub-categorias dependia ou não do idioma em que os informantes se expressaram. O nível de significância adotado foi de 0,05 e, sendo $m=5$ e $n=2$, ficou-se com 4 graus de liberdade. Neste caso, $\chi^2 = 9,49$. Estabeleceu-se como hipótese nula que, no nível Paradigmático, a ocorrência geral nas sub-categorias (Inadequação + Erro) independia da língua (português/castelhano) em que o informante foi solicitado a se manifestar. Por hipótese alternativa considerou-se que existia dependência entre a ocorrência das sub-categorias e a língua utilizada.

Ao ter sido obtido no nível Paradigmático o valor de 22,49 para χ^2 , pôde-se concluir que a hipótese alternativa foi validada, ou seja, dependia da língua falada pelo sujeito a ocorrência das sub-categorias consideradas, tendendo a haver maior inci-

dência de desvios quando os sujeitos se expressaram em português do que quando o fizeram em castelhano.

Efetou-se outro cálculo de X^2 nas mesmas condições de comparação, e com as hipóteses acima estabelecidas, agora para o nível Sintagmático. Também foi obtido resultado significativo. Verificou-se, portanto, como provavelmente verdadeira a hipótese alternativa, isto é, a ocorrência geral nas sub-categorias dependia da língua que o sujeito utilizava. Constatou-se, neste nível, um valor de 10,26, quando o valor crítico de comparação foi de $X^2=9,49$.

Aplicou-se um segundo X^2 mais específico, englobando as sub-categorias gerais Com Interferência (I-2, E-2A e E-2B) e Sem Interferência (I-1 e E-1), tanto no nível Paradigmático quanto no Sintagmático. A hipótese nula estabelecida foi: a ocorrência geral na classificação Com Interferência e Sem Interferência independia da língua em que o texto fora produzido. A hipótese alternativa considerava que a ocorrência geral dependia da língua utilizada. O nível de significância adotado para este caso foi também de 0,05 e para o grau de liberdade ($m=2$ $n=2$) encontrou-se 1 grau. O valor crítico detectado foi $X^2=3,84$.

Obteve-se $X^2=0,2$ para o nível Paradigmático e 0,5 para o Sintagmático. Portanto, a hipótese nula não foi rejeitada em nenhum dos níveis. Verificou-se que independia da língua a ocorrência da resposta Com e Sem Interferência. O resultado sugere, portanto, que os sujeitos cometeram desvios Com e Sem Interferência em ambas as línguas, não tendo sido relevante a língua em que se produziram tais emissões.

A Correlação de Postos (Siegel, 1956) foi o seguinte teste aplicado, agora inter-grupo. Procurou-se verificar se o bom desempenho numa língua correspondia ao bom desempenho na outra. Trabalhou-se com o nível de significância 0,05 e considerando que $n=6$, obteve-se o valor crítico para comparação de $r_c=0,71$ (Fisher e Yates, 1971). Foram comparadas as realizações dos sujeitos, no seu desempenho em língua portuguesa e castelhana, tanto no nível Paradigmático quanto Sintagmático, para as quatro categorias de respostas, a saber, Inadequada (I-1 e I-2), Errada (E-1, E-2A e E-2B), Sem Interferência (I-1 e E-1) e Com Interferência (I-2, E-2A e E-2B). Os resultados apresentaram-se globalmente significativos para G2 e não significativos para G1. G1 só obteve resultado significativo para a Resposta Sem Interferência (PI-1 + PE-1). G2 obteve resultados significativos na sub-categoria citada para G1 e em todas as outras sub-categorias do nível Sintagmático, ressalva feita à Resposta Com Interferência (SI-2 + SE-2A + SE-2B) que não atingiu, por uma margem mínima, o nível de significância estabelecido ($r=0,70$; $r_c=0,71$). Os dados sugerem que G1 utiliza as duas línguas de maneira independente, isto é, como dois desempenhos diferentes. Os resultados apresentados por G2 sugerem que este grupo se caracteriza por manter uma maior dependência entre as línguas, ao expressar-se em ambas. A transferência de elementos de uma língua para outra, em especial no nível Sintagmático, é mais relevante do que em G1.

Numa quarta etapa submeteu-se o material ao teste T de Wilcoxon, apropriado para a manipulação de amostras dependentes. O teste foi conduzido ao nível de significância de 0,05 e N mínimo =6, sendo o T_c ou termo de comparação igual a zero. Este teste, intra-grupo, permitiu comparar um tipo de desempenho com outro tipo de desempenho desse mesmo sujeito. Pelos resultados obtidos, os sujeitos tiveram estatisticamente o mesmo desempenho numa e noutra língua.

O quinto e último teste aplicado foi o teste U de Mann-Whitney (Siegel, 1956). Determinou-se também como nível de significância 0,05, sendo tanto N_1 como $N_2=6$. As sub-categorias de análise discriminadas no cálculo foram: PI-1, PI-2, PE-1, PE-2A, PE-2B, e as agrupadas: Sem Interferência (PI-1 + PE-1) e Com Interferência (PI-2 + PE-2A + PE-2B). Estabeleceu-se por hipótese nula para todas as sub-categorias de análise, que G1 teria desempenho equivalente a G2. Por hipótese alternativa aceitou-se que G1 teria um desempenho com menor incidência das categorias Erro e Inadequação, quer fossem Com ou Sem Interferência. Isto, considerando que seus anos de vivência no país deveriam

constituir variável mais relevante para o primeiro grupo do que para o segundo.

Os resultados inferidos deste teste informam que a variável tempo de permanência no país só foi significativa para G1 e no que concerne à transferência de letras do português para o castelhano (PE-2A). No relativo a todas as outras subcategorias e aos desvios Com e Sem Interferência, nos dois níveis, Paradigmático e Sintagmático, e em ambas as línguas, a produção verbal escrita destes imigrantes mostrou-se equivalente e independeu dos 10 anos de diferença cronológica que os se para quanto à sua radicação no país hospedeiro, no momento em que produziram os textos.

V-CONCLUSÕES

Os resultados obtidos, resumidos no item anterior em termos qualitativos e quantitativos, revelaram algumas tendências e particularidades do material analisado. Procurar-se-á nesta parte comentar o que foi constatado.

O primeiro χ^2 informou que os sujeitos tenderam a errar mais numa língua do que na outra, isto é, produziram mais desvios no português do que no castelhano. Este fato justifica-se se forem levadas em conta as circunstâncias lingüísticas dos informantes desta pesquisa. Todos eles adquiriram o segundo idioma quando já eram adultos e não seguiram estudos sistemáticos de língua portuguesa que lhes proporcionasse capacitação oral e escrita adequada. A aprendizagem do idioma sem um suporte didático específico tende geralmente a cristalizar, no comportamento verbal do indivíduo, emissões distorcidas de vários tipos. Este é o caso, em geral, dos imigrantes estabelecidos em São Paulo. Além disso, os modelos de aprendizagem a que estiveram expostos, possivelmente por tratar-se de situações menos formais, podem não se terem constituído nos melhores para o aprendizado em pauta. Considerando-se, por um lado, a importância da exposição a modelos na aprendizagem de línguas e, por outro lado, o fato de que, no uso diário, a maioria das vezes as pessoas não estão atentas ao padrão lingüístico que estão empregando, era de se esperar que ocorressem dificuldades nesta aprendizagem (Bandura, 1977; Staats, 1968), bem como transferências inadequadas do aprendido de uma para outra língua.

Os resultados do segundo χ^2 validaram a hipótese nula, isto é, a de que a ocorrência das categorias não estava relacionada com a língua em que os textos foram redigidos. Infere-se, portanto, que os informantes transferiram sua produção desviada do português para o castelhano e vice-versa. A análise qualitativa ratificou este resultado pois encontrou-se grande similaridade entre os desvios proativos, retroativos e intralinguais encontrados em ambas as línguas, para cada categoria e nível. Estes dados mostram a importância que teria para os sujeitos desta pesquisa a aprendizagem sistemática do português, a fim de obterem um desempenho normativamente adequado nos dois idiomas. Os efeitos resultantes do desconhecimento ou da não aplicação de uma regra idêntica à do idioma materno (cf. concordância) refletiram-se em ambas as línguas. Este efeito constatou-se também no que tange às regras da língua materna contrapostas às equivalentes, mas não idênticas, da língua adquirida posteriormente, no caso, o português (cf. preposição funcional de complemento direto e indireto). A falta de estabilidade na expressão destas regras atingiu os dois idiomas. Esta constatação é particularmente importante para o ensino de línguas. A tradição da análise contrastiva tem sempre enfatizado a questão da transferência de formas e estruturas da primeira para a segunda língua. Parece que alguma atenção deva também ser dada à transferência da segunda para a primeira.

Os resultados do teste da Correlação de Postos sugerem que G1 conseguiu separar os dois códigos com maior eficiência no seu comportamento verbal do que G2. O primeiro grupo só obteve resultado significativo quando não se tratava de bilingüismo mas de desconhecimento da língua em que se estava expressando; G2 também forneceu correlação significativa para esta resposta. Qualitativamente tais emissões estão vinculadas a problemas de expressão em língua escrita e dificuldades ortográfico-fonológicas, es

pecialmente no português. G2 apresentou correlação significativa em todas as subdivisões de categorias no nível Sintagmático (salvo na Resposta Com Interferência que, como já foi dito, não atingiu o nível por uma margem de 0,01).

Pode-se supor que no momento em que se produziram os textos, G2 já tivesse superado a fase de forte influência proativa e retroativa no nível Paradigmático, e se encontrasse no processo de isolar os dois idiomas no nível Sintagmático. G1, por sua vez, teria transposto esta etapa anteriormente. No entanto, este último grupo encontrou igualmente dificuldades nas Respostas Sem Interferência. Por conseguinte, o tempo não parece ter contribuído como fator de equilíbrio no que concerne a este tipo de dificuldade em nenhum grupo. A solução parece, pois, depender grandemente do apoio didático, sugerindo mais uma vez a relevância do ensino formal no domínio da língua.

No último teste aplicado, a hipótese nula foi validada, ou seja, a variável tempo de permanência no país, onde o convívio contínuo com a língua poderia exercer pressões no desempenho lingüístico do imigrante, não foi significativa. Fazem-se necessárias outras pesquisas que controlem características como nível de formação escolar, idade de aquisição do segundo idioma e outras, as quais permitirão conhecer variáveis mais relevantes que incidem na produção desviada destes imigrantes.

O citado teste só ofereceu um único resultado significativo: o grupo 1 transferiu um número relevante de lexias do português para o castelhano. Trabalhos nesta linha têm demonstrado que o uso intenso e maiores oportunidades de comunicação são variáveis significativas no domínio de uma língua sobre outra (Hornby, 1977). Também pesquisas sobre retenção de memória verificaram que estímulos paralelos semelhantes favorecem o declínio da retenção mnemônica (Overbeke, 1976). Dois fatores participam relevantemente no processo: a primazia e a recentidade. Nos estímulos, os primeiros e os últimos são os mais retidos pelos sujeitos. Na memória imediata é o efeito de recentidade que entra em jogo, na memória a longo prazo, o de primazia.

Portanto, o tempo por si só não parece ser um fator preponderante, mas influente e imbricado no processo. Considerando-se que a) ele representa a linha cronológica em que se produz a atuação verbal do indivíduo; b) que a vivência lingüística nesse tempo privilegia o português, não só em termos de frequência mas de uso social; e c) que os dois sistemas do bilingüe são muito próximos (estímulos semelhantes), é razoável que o grupo de informantes que sofreu a exposição mais intensa à língua de convívio tenda a transferir uma quantidade maior de lexias para o castelhano do que para o português.

Em várias ocasiões, na análise do material, o nível Sintagmático sobressaiu com relação ao Paradigmático. Este fato leva a pensar que os dois eixos da linguagem demandam custos diferentes em termos de produção lingüística por parte dos informantes. É possível que o nível onde entrem em jogo, prioritariamente, as relações funcionais e semânticas entre lexias seja mais crítico para os sujeitos do que o Paradigmático que se atém preferentemente à unidade mínima de comportamento. Dito de outra maneira, pode ser que o procedimento mental de escolha de uma unidade lexical e aquele que compromete mais de uma unidade exijam competências e hierarquias diferentes. Talvez a escolha lexical privilegie a semântica (necessidade do indivíduo de veicular conteúdos) e a distribuição do léxico requeira prioritariamente a manifestação do conhecimento gramatical do idioma, isto é, as regras de relação entre lexias.

Esta dicotomia apesar da sobreposição que a caracteriza, demonstrou ser um mecanismo útil e didático na classificação, detecção e especificação das emissões encontradas no corpus. E sugere, ainda, que em línguas de proximidade homonímica como o português e o castelhano, as influências proativas e retroativas ocorrem igualmente no léxico e na gramática, com uma proporção relevante nesta última.

O maior número de emissões desviadas em ambas as línguas e grupos concentrou-se na categoria Erro. O Erro Sem Interferência, no nível Sintagmático, apresentou desvios similares em ambas as línguas, basicamente de estruturação e organização formal deficientes do material lingüístico. No nível Paradigmático, as respostas desviadas mais nume-

rosas foram devidas à ignorância ou à assimilação equivocada de regras específicas de cada idioma.

O Erro Com Interferência por Empréstimo foi quantitativamente preferido pelos sujeitos comparado ao Hibridismo. O procedimento de Empréstimo sugere um esforço verbal menor por parte do indivíduo do que o Hibridismo. No primeiro o bilingüe lança mão de um dos códigos de que dispõe, o da LP. No Hibridismo, o comportamento cognitivo envolvido parece ser mais complexo (escolha da forma de uma das línguas e significado da outra, adaptação da forma ou do conteúdo de uma à outra, ou, ainda, associação da forma ou do conteúdo de uma com outro elemento não equivalente da outra). Estes últimos fenômenos constataram-se não só na lexia, mas no interior das zonas lexicais e nos segmentos frasais. O ato de mescla supõe dois códigos em conflito e alguma manipulação das diferenças no ato de fala. O Empréstimo, ao contrário, implica numa opção única, provavelmente de menor custo.

Em língua castelhana, G2 destacou-se sensivelmente de G1 ao fornecer porcentuais superiores de respostas híbridas. Este resultado pode ser atribuído à pressão exercida pelo português sobre o castelhano, desestabilizando a expressão neste último. Há lógica em supor que o Hibridismo se apresente em diversos graus qualitativos e quantitativos conforme a distância tipológica entre as línguas em convívio. No entanto, com relação aos idiomas em análise pode-se observar, pelas emissões híbridas do corpus, uma grande permeabilidade entre as duas línguas. As respostas híbridas encontradas refletem a coincidência de aspectos lexicais, morfo-sintáticos, ortográficos e/ou fonológicos de natureza formal e/ou semântica dos dois idiomas num mesmo elemento. A amplitude homônima entre as línguas em questão parece não oferecer muitas restrições a tal procedimento.

O número elevado de Empréstimos no nível Paradigmático de lexias gramaticais castelhanas detectadas no corpus em português é um dado que merece atenção. Estas lexias caracterizaram-se pela grande semelhança formal e total identidade semântica com as correspondentes portuguesas. Se a memória de recentidade vem apoiar algumas das tendências que afloraram nesta pesquisa, a memória de primazia parece estar presente na produção destas lexias gramaticais. Este fenômeno, constatado especialmente mas não exclusivamente em PE-2A, pode estar associado ao processo psicolinguístico denominado por Selinker (1974) fossilização.

As reflexões que o corpus propicia nesta direção abrem perspectivas para a pesquisa e o ensino. Estudos que verificassem a correlação entre a atuação da memória de recentidade e a de primazia no que tange à produção qualitativa de desvios, permitiriam controlar e reduzir este fenômeno, pela utilização de material didático específico e apropriado, bem como o desenvolvimento e teste de tecnologias de ensino mais eficientes.

Os dados qualitativos sugerem também que os informantes preferiram a subdiferenciação à superdiferenciação, comportamento que pode estar relacionado com a semelhança entre as línguas e com o princípio de economia de resposta (Staats, 1968; Overbeke, 1976; Ben-Zeew, 1977).

Pelo exposto pode concluir-se que, ao imigrante castelhano, não basta a convivência social com a língua portuguesa para manejá-la com a correção normativa adequada. Também pouco é suficiente, para manter a sua língua materna como código independente, ter adquirido e utilizado o castelhano até a idade adulta. A aprendizagem do português como um sistema linguístico particular parece ser um caminho seguro não só para a integração social, mas para a manutenção de sua primeira língua, sendo igualmente relevante, neste último caso, dispor de oportunidades de uso formal para tal manutenção.

Estas observações solidificam a idéia de que trabalhos de comparação entre os dois idiomas, a partir do levantamento e análise dos desvios cometidos por bilingües e aprendizes, possibilitarão a elaboração de material didático eficaz para o ensino do português e do castelhano a hispano-falantes e luso-falantes respectivamente. As características peculiares destas duas línguas em contato num só indivíduo requerem tratamento especial no que tange à metodologia de aprendizagem.

O castelhano e o português, cada vez mais, são línguas que se irmanam, impelidas pelas circunstâncias geográficas, econômicas, sociais e culturais das comunidades que os utilizam. A atenção que estão merecendo, reflexo das necessidades deste fim de século, deve sensibilizar os educadores e incentivá-los a realizar um trabalho eficaz. Esta pesquisa espera ser um estímulo neste sentido.

VI- BIBLIOGRAFIA

- BANDURA, A. (1977) Social learning theory. New Jersey, Prentice-Hall, vol. 1-2.
- BEN-ZEEW, S. (1977) Mechanisms by which childhood bilingualism affects understanding of language and cognitive structures. In HORNBY, P.A. (vide infra)
- BONTEMPO, E. (1975) Observação: um método para estudo do comportamento. In WITTER, G.P. Ciência, ensino e aprendizagem. São Paulo, Alfa-Omega.
- FISHER, R.A & YATES, F. (1971) Tabelas de estatística: para pesquisa em Biologia, Medicina e Agricultura. São Paulo, Polígono/EDUSP.
- HORNBY, P. e outros (1977) Bilingualism. Psychological, social and educational implications. New York, Academic Press, Inc.
- LADO, R. (1973) Lingüística contrastiva, lenguas y culturas. (Trad. por Joseph A. Fernández). Madrid, Alcalá.
- OVERBEKE, M.V. (1976) Mécanismes de l'interférence linguistique. Madrid, Fragua.
- PEABODY (1968) Peabody language development kits. American Guidance Service Inc., Publisher's Building, Circle Pines, Minn 55014, Level P., X-13.
- POTTIER, B. (1977) Lingüística general. (Versión española de María Victoria Catalina) Madrid, Gredos.
- RICHARDS, J.C. (ed.) (1974) Error analysis. Perspectives on second language acquisition. London, Longman.
- SELINKER, L. (1974) Interlanguage. In RICHARDS, J.C (ed.) (vide supra)
- SIEGEL, S. (1956) Non-parametric statistics for the behavioural sciences. New York, MacGraw Hill.
- STAATS, A.W. (1968) Learning, language and cognition. New York, Holt, Rinehart and Winston Inc.
