

IDENTIDADE DE REPRESENTAÇÕES BÁSICAS E FATORES
DE DIFERENCIAÇÃO SUPERFICIAL NO COMPONENTE
FONOLÓGICO DE LÍNGUAS COGNATAS

por

MARIA BERNADETE MARQUES ABAURRE

Dissertação apresentada ao
Departamento de Linguística do
Instituto de Filosofia e Ciências
Humanas da Universidade
Estadual de Campinas como requi-
sito parcial para obtenção do
grau de Mestre em Linguística

AB181

Campinas

1973

b18i

15/BC

A pesquisa objeto desta dissertação foi possibilitada, em parte, por uma bolsa de pós-graduação concedida pelo Programa Unificado de Pós-Graduação em Linguística da UFRJ com recursos da Fundação Ford (Grant nº 67-450 A).

LISTA DE SÍMBOLOS

- [] - representação fonética ou conjunto de propriedades.
- | | - representação fonológica (forma básica, representação subjacente)
- + - junção de morfema
- ≡ - junção de palavra
- └ - forma verbal
- V - vogal
- C - consoante
- G - glide
- R - ressonante
- [t', d'] - [t, d] palatalizados
- ∅ - "zero"

SUMÁRIO

Introdução.....	p. 6
1 - O Sistema Sonoro do Português: Segmentos Subjacentes.....	p. 13
1.1 Considerações preliminares.....	p. 13
1.1.1 Níveis linguisticamente significativos....	p. 13
1.1.2 "Fonema" e traços distintivos.....	p. 14
1.1.3 A análise do componente fonológico: representações lexicais, fonológicas e fonéticas.....	p. 15
1.1.3.1 Representações fonéticas.....	p. 17
1.1.3.2 Representações lexicais e representações fonológicas.....	p. 18
1.2 Matrizes de traços distintivos para os segmentos do português.....	p. 24
1.2.1 Matriz das obstruintes e ressonantes.....	p. 27
1.2.2 Matriz das vogais.....	p. 29
1.3 Conclusão.....	p. 30
2 - Os Estados Lexicais e o Âmbito de Aplicação das Regras Fonológicas.....	p. 32
2.1 Considerações preliminares.....	p. 32
2.2 O reconhecimento de estratos lexicais no português.....	p. 35
2.3 As propriedades diacríticas.....	p. 37
2.4 Sincronia e diacronia.....	p. 42
2.5 Conclusão.....	p. 47
3 - As Formas do Presente do Indicativo no Espanhol e no Português: Evidências para Identidade de Representações Básicas.....	p. 49
3.1 Considerações preliminares.....	p. 49
3.2 A classe verbal no português.....	p. 51
3.2.1 As noções gramaticais de número, pessoa, tempo, aspecto, modo e voz.....	p. 53
3.2.1.1 Número.....	p. 54
3.2.1.2 Pessoa.....	p. 54
3.2.1.3 Tempo.....	p. 55
3.2.1.4 Aspecto.....	p. 56
3.2.1.5 Modo.....	p. 56

3.2.1.6 Voz.....	p. 57
3.3 Dados considerados.....	p. 60
3.4 Análise dos dados.....	p. 65
3.4.1 As formas do presente do indicativo.....	p. 65
3.4.1.1 Primeira conjugação.....	p. 65
3.4.1.2 Segunda conjugação.....	p. 89
3.4.1.3 Terceira conjugação.....	p. 91
3.4.2 Derivação das formas.....	p. 97
3.4.2.1 Presente do indicativo: 1a. conjugação regular.....	p. 98
3.4.2.2 Presente do indicativo: 2a. conjugação regular.....	p. 100
3.4.2.3 Presente do indicativo: 3a. conjugação regular.....	p. 102
Conclusão.....	p. 104
Notas.....	p. 107
Introdução.....	p. 107
Capítulo 1.....	p. 111
Capítulo 2.....	p. 138
Capítulo 3.....	p. 143
Lista de regras.....	p. 168
Bibliografia.....	p. 171

INTRODUÇÃO

Com base nos princípios teóricos da fonologia gerativa apresentamos, neste trabalho, uma análise comparativa de algumas formas verbais do espanhol e do português.¹ Estaremos considerando as formas do presente do indicativo nas três conjugações regulares das duas línguas, e o objetivo principal do estudo é examinar a hipótese segundo a qual podem ser estabelecidas, na maioria dos casos, as mesmas representações subjacentes para o espanhol e para o português. Examinaremos também a hipótese de que as diferenças nas manifestações fonéticas das duas línguas resultam quase sempre da aplicação de maior número de regras no caso do português.

Não será nosso objetivo apresentar aqui um tratamento exaustivo do assunto. Procuraremos essencialmente demonstrar como pode ser elaborado um trabalho nesta linha, e indicar as possíveis implicações advindas de uma tal comparação para o melhor entendimento da natureza e grau de diferenciação entre as duas línguas.

Limitamo-nos à consideração das formas do presente do indicativo principalmente devido ao pouco tempo de que dispusemos para examinar os dados. Lembremos, no entanto, que o estudo comparativo das demais formas pode ser conduzido na mesma linha, tanto para os verbos regulares como para os considerados irregulares (i.e., naturalmente, para os itens pertencentes a outras classes gramaticais). Da consideração de maior número de dados, aliás, é mesmo provável que surjam novos índices de diferenciação entre o português e o espanhol, i. e.,

pode ser que haja outras explicações para as diferenças nas manifestações superficiais (além da que já foi mencionada anteriormente, maior número de regras em um dos casos). Algumas possibilidades seriam, por exemplo: diferença na ordenação de um mesmo conjunto de regras, supressão de alguma(s) regra(s), etc.

A análise na qual nos baseamos para o espanhol é a proposta por Harris (1969). Faremos, portanto, freqüentes referências a esta obra nos capítulos seguintes.

Além do objetivo básico de comparar as formas nas duas línguas e tentar explicar a diferenciação no âmbito fonético, há alguns aspectos que procuraremos enfatizar ao considerar as formas do português. Tais aspectos dizem respeito a pressupostos teóricos nos quais nos baseamos. Assim é que as regras apresentadas procuram explicitar algumas das alternâncias morfofonêmicas responsáveis pela configuração do padrão sonoro da língua, depreensíveis a partir do exame destas formas verbais, apenas. Procuraremos também mostrar como pode ser tratado o problema da sistematização dos processos flexionais operantes nesta classe de palavras. Por serem geralmente da mesma natureza os processos morfofonêmicos que se manifestam nas diversas classes gramaticais das línguas, não nos restringiremos aos verbos, buscando, sempre que possível, exemplos em outras classes, ora para provar a maior generalidade de determinado processo, ora para reforçar a motivação de determinada regra formulada. O trabalho deve ser também entendido, portanto, como uma tentativa de demonstração de como pode ser feito um estudo do componente fonológico de uma gramática ge-

rativa.² Não se trata, repetimos, de uma análise exaustiva deste componente da gramática do português, uma vez que trabalhamos com um número bastante limitado de dados. É necessário ter sempre em mente que a consideração futura de maior número de dados poderá levar a modificações nas regras aqui propostas, isto porque, quanto mais extenso o corpus considerado, maiores as possibilidades de se chegar a descobrir os processos fonológicos gerais que subjazem às manifestações fonéticas de determinada língua.

Ao organizar este trabalho consideramos também como um dos seus principais objetivos a demonstração de que a utilização deste modelo de análise, baseado nas mais modernas tendências lingüísticas, permite atingir uma maior generalização com relação aos fenômenos operantes na língua, ao mesmo tempo que estabelece, entre os diversos processos envolvidos, uma série de relações importantes. Estas relações, que constituem a própria essência de um sistema lingüístico, obscureciam-se, via de regra, nas descrições tradicionais encontradas na grande maioria das gramáticas escolares, ao serem tratados isoladamente problemas como "plural dos nomes em -ão", "plural dos nomes em -al, -el, -ul", "plural dos nomes em -il", "plural dos nomes em -r, -n, -z", etc. A simplicidade do processo envolvido, bem como a amplitude de sua aplicação na língua, escapavam freqüentemente aos pesquisadores que, influenciados muitas vezes por problemas secundários de ordem fonética, deixavam de ver o essencial, ou seja, os processos morfofonêmicos gerais que relacionam os diversos casos. A gramática da língua, resultante deste tipo tradicional de análise, apresenta

va uma complexidade excessiva e consistia em um agrupamento muitas vezes arbitrário de fatos desconexos. Não se podia esperar unicidade e simplicidade de uma gramática assim construída. De acordo com a teoria transformacional, na qual nos baseamos, os conceitos de simplicidade e economia estão intimamente relacionados ao de generalidade. Uma gramática é mais simples e as regras propostas são mais econômicas, na medida em que refletem processos mais gerais da língua. Esforçemo-nos, portanto, por descobrir o maior número possível de generalizações importantes dentro da pequena parte do componente fonológico do português que nos propusemos abordar.

Procuraremos ainda evidenciar a estreita relação entre a fonologia e a morfologia da língua. O tratamento isolado destes dois níveis, proposto pelo estruturalismo lingüístico e levado inclusive a extremos por alguns de seus adeptos, revela-se, com efeito, artificial. A teoria transformacional propõe que tanto os problemas fonológicos como os morfológicos sejam estudados dentro do componente fonológico da gramática, relacionado, por sua vez, com os componentes sintático e semântico.³

Ficará evidente, assim que introduzirmos, a partir do capítulo 3, as formas verbais propriamente ditas, que as manifestações fonéticas, responsáveis por tantas interpretações imprecisas em estudos tradicionais, nunca são utilizadas neste trabalho como ponto de partida para a análise. Elas são sempre entendidas como sendo o resultado final da aplicação de uma série de regras e formas abstratas. O estabelecimento destas formas não é de modo algum arbitrário, mas coerente, den-

tro de determinados critérios previamente estabelecidos, conforme se verá por ocasião de sua apresentação. A análise a partir de formas básicas fica enormemente simplificada porque impede-se, que certos detalhes fonéticos, restritos por vezes a determinadas variantes da língua, assumam uma importância excessiva. Estes detalhes que, em última análise, identificam também as diversas variantes de uma mesma língua, explicam-se, via de regra, por regras simples, aplicadas na superfície. Observa-se, pois, que é também simplificado, de acordo com esta maneira de considerar os fatos, o próprio estudo da variação lingüística. Enquanto as regras morfofonêmicas aplicam-se à língua em um nível mais profundo, as regras de detalhe fonético, sua natureza e/ou ordenação são responsáveis pela maior parte da variação. Assim, como se verifica no caso da comparação entre duas línguas distintas, uma variante X de uma língua A pode ter mais ou menos regras de superfície do que a variante Y da mesma língua, ou pode ter as mesmas regras da variante Z, ordenadas de modo diferente.⁴ As regras morfofonêmicas, no entanto, devem ser, na medida do possível, comuns a todas as variantes da língua (algumas delas podem inclusive ser as mesmas no caso de línguas de relação genética muito próxima como o português e o espanhol, conforme tentaremos demonstrar no decorrer da exposição). Não nos preocupamos em formular regras que dêem conta de todos os detalhes fonéticos das formas verbais consideradas. Achamos que isso já constituiria assunto para um outro trabalho que tratasse, por exemplo, extensivamente, da variação dialetal no português. Nesse caso seria indispensável levar em consideração todos os

detalhes fonéticos, uma vez que as diferenças entre os diversos dialetos de uma mesma língua manifestam-se, geralmente, na superfície. No entanto, embora a análise apresentada diga respeito essencialmente a um nível mais profundo (já que procuramos depreender regras que possam ser consideradas básicas para o português e para o espanhol e que tenham como âmbito de aplicação, nos dois casos, o que Hockett (1958) chama de "âmago comum da língua" (ingl. 'common core')), alguns detalhes fonéticos não podem, evidentemente, ser ignorados, pois constituem muitas vezes o fator de diferenciação superficial entre as duas línguas. Sempre que a menção de alguma dessas regras superficiais se faz necessária para o esclarecimento de determinados problemas, a variante do português na qual nos baseamos é a variante regional capixaba (falada em Vitória e provavelmente em grande parte do Estado do Espírito Santo). Os dados que utilizamos nesses casos refletem a nossa intuição de falante nativa desta variante.

As citações, por vezes extensas, que fazemos acompanhar certas discussões de ordem teórica, justificam-se pelo caráter de divulgação que gostaríamos também de imprimir a este trabalho. Em princípio, qualquer aplicação de uma teoria determinada dispensa a apresentação dos princípios básicos desta mesma teoria, pois parte-se do pressuposto de que tais princípios já são conhecidos. Obras de divulgação da teoria fonológica gerativa escritas para leitores de língua portuguesa são, entretanto, praticamente inexistentes até o momento e foi exatamente este o motivo que nos levou a procurar incluir em todos os capítulos, sempre que possível, informações teóri

cas indispensáveis à compreensão da análise.⁵ As informações básicas aqui fornecidas não dispensam, no entanto, uma certa familiaridade com os pressupostos teóricos envolvidos, bem como o conhecimento de algumas das principais obras sobre o assunto.

Introduziremos, no capítulo 1, algumas das noções teóricas que consideramos essenciais para a compreensão dos capítulos seguintes, bem como as matrizes de traços distintivos para os segmentos subjacentes do português. No capítulo 2 trataremos em certo detalhe do problema da estratificação lexical na língua, i. e., da necessidade de categorizar os itens que fazem parte do léxico conforme constituam ou não âmbito de aplicação para um certo conjunto de regras fonológicas. Finalmente, no capítulo 3, apresentaremos a comparação entre as formas verbais propriamente ditas do espanhol e do português, e as conclusões a que chegamos a partir de sua consideração.

CAPÍTULO 1

O SISTEMA SONORO DO PORTUGUÊS: SEGMENTOS SUBJACENTES

1.1 Considerações preliminares

Antes da apresentação das matrizes de traços distintivos para cada um dos segmentos que subjazem às manifestações fonéticas no português, faz-se necessário discutir os seguintes aspectos teóricos:

- (a) - níveis reconhecidos como lingüisticamente significativos pela teoria fonológica gerativa;
- (b) - status de fonema e dos traços distintivos nas análises do componente fonológico.

Passaremos a seguir à consideração destes tópicos.

1.1.1 Níveis lingüisticamente significativos

O modelo estrutural de análise lingüística distingue, em fonologia, três níveis significativos: o nível fonético, o nível fonêmico e o nível morfofonêmico. A unidade teórica mínima deste modelo é o fonema, de status relevante, em torno do qual gira toda a análise fonológica.

A fonologia gerativa distingue apenas dois níveis lingüisticamente significativos: o nível fonológico (ou nível de abstração) e o nível fonético (ou nível da realidade). O nível fonológico é o das representações subjacentes que constituem o input para a aplicação das regras fonológicas. O nível fonético é o das manifestações superficiais, output ou produto fonético final. Entre estes dois níveis não se reconhece nenhuma etapa significativa, correspondendo o primeiro à competência (ingl. "competence") do falante nativo com relação à própria língua, e o segundo à realização (ingl. 'performance') ou maneira como se atualizam, na cadeia da fala, os conhecimentos interiorizados que possui o falante sobre o componente

fonológico da gramática de sua língua.¹ Assim, um só segmento depressível no nível fonológico pode ter várias manifestações físicas no nível fonético, conforme o condicionamento que sobre ele exerçam os segmentos vizinhos, o grau de intensidade do acento; etc.

O reconhecimento dos níveis significativos da competência e da realização, em lingüística, foi um dos pontos de partida para o desenvolvimento da teoria transformacional. Esta distinção é essencial, portanto, para a consideração de cada um dos componentes da gramática (v. nota 2, introdução) e explica porque a análise se processa, também no componente fonológico, em dois níveis apenas, conforme mencionamos no início desta seção (nível fonológico ou da abstração e nível fonético ou da realidade).

1.1.2 "Fonema" e traços distintivos

Em torno do fonema, unidade teórica significativa, gira a análise fonológica proposta pela lingüística estrutural (v. l. 1.1). A partir da observação das manifestações fonéticas, identificam-se os fonemas ou segmentos opositivos das línguas, bem como suas variantes posicionais ou alofones. No caso dos fonemas a ocorrência é não condicionada, enquanto que as variantes posicionais são determinadas pela natureza do ambiente segmental e supra segmental. Uma vez identificados são os segmentos opositivos organizados em paradigmas e qualquer alternância fonológica ou morfológicamente condicionada entre segmentos previamente reconhecidos como fonemas opera-se no chamado nível morfofonêmico.

O fonema é, portanto, neste modelo de descrição, o centro

de toda a análise fonológica.

A teoria transformacional, por outro lado, propõe que a análise do componente fonológico se faça a partir de traços distintivos.² A unidade teórica mínima passa então a ser, na fonologia gerativa, o traço distintivo, e não mais o fonema. Este termo continua a ser empregado em alguns trabalhos, no entanto, mas como mera convenção, isto é, quando se quer fazer referência a uma determinada combinação de traços distintivos em uma língua qualquer. Costuma-se então falar em fone ma sistemático. Evitaremos o termo, nesta análise do português, por motivos já aqui discutidos (v. nota 1 deste capítulo). Sempre que, no decorrer deste trabalho, for necessário mencionar algum conjunto de traços distintivos no português, utilizaremos segmento, termo a nosso ver mais neutro, para facilidade de exposição. Os símbolos fonéticos empregados devem também ser entendidos como abreviações gráficas de complexos de traços nas matrizes das representações lexicais, fonológicas e fonéticas de que falaremos a seguir.

1.1.3 A análise do componente fonológico: representações lexicais, fonológicas e fonéticas.

Algumas considerações sobre o componente fonológico de uma gramática gerativa foram já feitas na introdução (nota 2).

Abordaremos o mesmo assunto nesta seção para que possam ser introduzidas mais algumas informações teóricas indispensáveis ao entendimento das representações que constituem o input e output da análise fonológica.³

Vimos já que uma gramática é constituída de três componen

tes: sintático, semântico e fonológico. As relações entre eles são assim descritas por Chomsky e Halle (1968):

(...) a grammar contains a syntactic component, which is a finite system of rules generating an infinite number of syntactic descriptions of sentences. Each such syntactic description contains a deep structure and a surface structure that is partially determined by the deep structure that underlies it. The semantic component of the grammar is a system of rules that assigns a semantic interpretation to each syntactic description, making essential reference to the deep structure and possibly taking into account certain aspects of surface structure as well. The phonological component of the grammar assigns a phonetic interpretation to the syntactic description, making reference only to properties of the surface structure, so far as we know. The structural description assigned to a sentence by the grammar consists of its full syntactic description, as well as the associated semantic and phonetic representations. Thus the grammar generates an infinite number of sentences, each of which has a phonetic and semantic interpretation: it defines an infinite sound-meaning correspondence, this correspondence being mediated by the abstract syntactic component and the structures it generates. (6-7)

As estruturas superficiais, output do componente sintático, constituem-se no input do componente fonológico. Isto quer dizer que são, ao mesmo tempo, resultado da aplicação de regras sintáticas independentemente motivadas e ponto de partida para a aplicação de regras fonológicas, responsáveis pelas interpretações fonéticas finais. É, portanto, a partir das estruturas sintáticas superficiais, que se realiza a análise do componente fonológico de uma língua.⁴

Trataremos a seguir das representações fonéticas, lexicais e fonológicas e da maneira como se processa a análise fonológica propriamente dita.

1.1.3.1 Representações fonéticas

As propriedades essenciais das línguas naturais constituem os chamados "universais lingüísticos".⁵ Uma das preocupações da lingüística moderna é exatamente a descoberta destes universais que representam o que se convencionou chamar de "gramática universal", e com base nos quais é possível avaliar o grau de complexidade das gramáticas específicas. Embora entrem também em jogo outros fatores, uma gramática pode ser considerada tanto mais complexa quanto maior for o número de regras específicas que apresente.

Quanto aos universais fonéticos substantivos, alguns seriam:

- (a) - a existência de um conjunto universal de traços fonéticos;
- (b) - a possibilidade de combinação destes traços em segmentos;
- (c) - a possibilidade de combinação destes segmentos (complexos de traços) em enunciados;
- (d) - a existência de restrições específicas quanto às combinações previstas em (a) (simultâneas) e em (b) (seqüenciais).

Com base no conjunto universal de traços fonéticos e em suas possibilidades de combinação pode-se postular ainda a existência do que seria a classe das representações fonéticas possíveis. Todas as manifestações fonéticas, em todas as línguas naturais, seriam tiradas desta classe de representações fonéticas possíveis.

Conclui-se então que a representação fonética de qualquer enunciado em qualquer língua é uma matriz em que as colunas ho

rizontais representam traços fonéticos e as colunas verticais, segmentos ou combinações de traços.⁶

Quando se trabalha com o componente fonológico de uma língua, o produto final da análise é justamente o conjunto das representações fonéticas desta língua.

1.1.3.2 Representações lexicais e representações fonológicas

Uma vez estabelecido serem as estruturas superficiais o input para a análise fonológica, resta considerar a própria constituição destas estruturas.

Cada estrutura superficial é gerada pelo componente sintático da gramática e consiste em uma cadeia de elementos mínimos, denominados "formativos". Os formativos são categorizados de acordo com as funções sintáticas que podem exercer e quanto às propriedades semânticas que apresentam. As categorias a que pertencem determinam também as suas representações subjacentes, de natureza abstrata, que comentaremos oportunamente.

Para exemplificar a maneira pela qual um formativo é categorizado, pode-se pensar, por exemplo, em homem, que pertence, entre outras, às seguintes categorias: "nome", "animado", "macho", etc. No léxico, parte do componente sintático da gramática, são indicadas todas as categorias a que pertencem os vários formativos e que podem também ser entendidas como traços destes formativos (que apresentam, portanto, traços fonológicos, sintáticos e semânticos). Partindo-se do princípio de que as entradas lexicais dos formativos não passam de listas das categorias às quais eles pertencem e de que as es-

estruturas superficiais são constituídas de cadeias de formativos, subentende-se que nestas estruturas superficiais estejam representadas de maneira completa as categorizações de cada formativo. Assim, o input para a análise fonológica caracteriza-se já, como output sintático, por traços fonológicos, sintáticos e semânticos.

Por considerarmos de importância para a compreensão do processamento da análise fonológica, transcrevemos aqui o trecho em que Chomsky e Halle (1968) abordam o problema da subdivisão em "frases" das cadeias de formativos que constituem as estruturas superficiais:

The surface structure must indicate how the string of formatives it contains is subdivided into "phrases", each phrase being a certain continuous substring of the string of formatives. The analysis of strings into phrases is a "proper bracketing" in the sense that phrases can overlap only if this is contained in the other. Thus, if A, B, C are formatives, the surface structure of the string ABC cannot specify AB as a phrase and BC as a phrase, for the string may be bracketed either as ((AB)C) or as (A(BC)) but not in both ways simultaneously (7).

Prosseguimos com a citação para que se tenha uma idéia da diagramação prevista pela teoria:

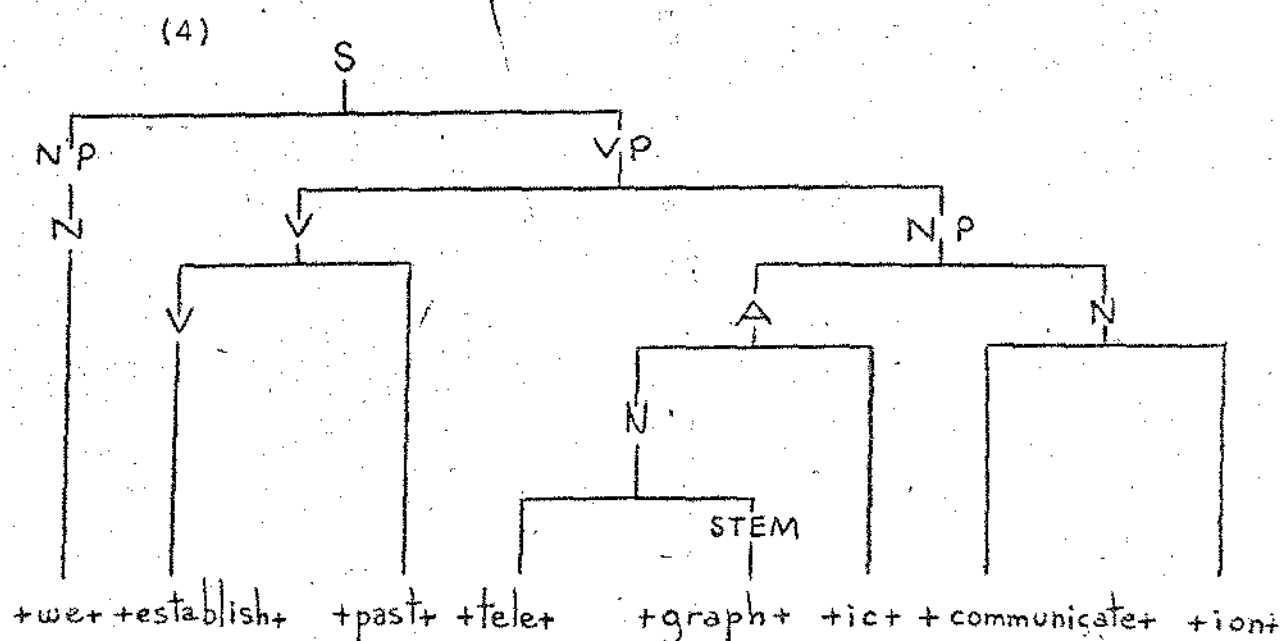
The phrases furthermore are assigned to certain categories, and this information may be represented by putting labels on the brackets. Take, for example, the sentence (3):

(3) we established telegraphic communication

In (3) the string underlying we is assigned to the same category as the string underlying telegraphic communication, namely, to the category "noun phrase". Similarly, the other phrases are assigned to certain universal categories. We will make the empirical assumption that the surface structure of a sentence is precisely a proper bracketing of a string of formatives, with the bracketed substrings (the phrases) assigned

to categories selected from a certain fixed universal set of categories. The complete string is assigned to the category "sentence" (S); the other phrases are also assigned to categories that are provided by general linguistic theory, such as the categories "noun phrases" (NP) and "verb phrase" (VP). These universal categories are on a par with the phonetic categories (bilabial closure, frontness, etc.) provided by universal phonetic theory. As we noted earlier, the categories of universal phonetic theory determine a certain infinite class of possible phonetic representations from which the phonetic forms of sentences of any human language are drawn. Similarly, the universal set of phrase categories (NP, VP, etc.), together with the universal lexical categories (noun, verb, adjective) and the universal lexical features that define the class of "possible formatives", provide us with an infinite class of possible surface structures, from which the surface structures of sentences of any particular language are drawn. In other words, general linguistics should provide definitions, in terms independent of any particular language, for the notions "possible phonetic representation" and "possible surface-structure". The grammar of each language relates phonetic representations to surface structures in a specific way; and, furthermore, it relates surface structures to deep structures (grifo nosso), and, indirectly, to semantic interpretations, in ways that are beyond the scope of our present study.

To give a concrete example, the grammar of English might assign to the sentence (3) a surface structure which can be represented in the equivalent forms (4) and (5):



(5) [S[NP[N+we+]N]NP] [VP[V[V+establish+]V+past+]V[NP[A[N+tele+][STEM+graph+]STEM]N+ic+]A[N[V+communicate+]V+ion+]N]NP]VP]S.

The interpretation of the notational devices used in (4) and (5) should be obvious. We intend these representations to indicate that the formative we is both an N and an NP, the formative establish a V, the formative string tele graph ic communicate ion an NP, the full string an S, etc. Furthermore, each formative has an analysis as a set of intersecting categories, in a way that we shall specify in more detail below. The symbols represent formative boundaries which, by convention, automatically mark the beginning and end of each formative. (7-9)

Conhecidas a origem e a constituição das estruturas superficiais, pode-se passar à consideração das representações lexicais e fonológicas. Convém inicialmente lembrar a conceituação de representações fonéticas, apresentada em 1.1.3.1. Vimos que uma representação fonética é uma matriz de traços resultante da aplicação, a uma estrutura superficial, do sistema de regras que constituem o componente fonológico da gramática. Como as estruturas superficiais são cadeias de formativos, tem-se que as regras fonológicas se aplicam, em última análise, aos formativos dessas estruturas. Os formativos, que podem ser lexicais (cf. no diagrama proposto por Chomsky e Halle: we, establish, tele, graph, communicate) e gramaticais (past, ic, tion), representam-se também por matrizes de traços (sintáticos, semânticos e fonológicos) que vêm a ser justamente as representações lexicais e fonológicas.

Há casos em que as regras fonológicas se aplicam diretamente aos formativos conforme apareçam nas entradas lexicais.⁸ Tais formativos, extraídos diretamente do léxico, constituem

as chamadas representações lexicais. Nem sempre, no entanto, as estruturas superficiais, output do componente sintático, representam o input ideal para o componente fonológico. Para obviar eventuais discrepâncias a gramática prevê uma série de regras de reajuste (ingl. 'readjustment rules') que, aplicadas a estas estruturas superficiais ("inadequadas" como input fonológico), preparam seus formativos para a análise fonológica. As representações resultantes da aplicação de regras de reajuste são as representações fonológicas.

Vejamos como operam as regras de reajuste:

(...) the surface structure must meet two independent conditions: first, it must be appropriate for the rules of phonological interpretation; second, it must be "syntactically motivated", that is, it must result from the application of independently motivated syntactic rules. Thus we have two concepts of surface structure: input to the phonological component and output of the syntactic component. It is an empirical question whether these two concepts coincide. In fact, they do coincide to a very significant degree, but there are also certain discrepancies. These discrepancies (...) indicate that the grammar must contain certain rules converting the surface structures generated by the syntactic component into a form appropriate for use by the phonological component. In particular, if a linguistic expression reaches a certain level of complexity, it will be divided into successive parts that we will call "phonological phrases", each of which is a maximal domain for phonological processes (grifo nosso). In simple cases the whole sentence is a single phonological phrase; in more complex cases the sentence may be reanalyzed as a sequence of phonological phrases. The analysis into phonological phrases depends in part on syntactic structure, but it is not always syntactically motivated in the sense just mentioned. If the syntactic component were to be connected to an orthographic rather than a phonetic output system, the reanalysis into phonological phrases would be unnecessary. Writers, unlike speakers, do not run out of breath, and are not subject

to other physiological constraints on output that require an analysis into phonological phrases.

In addition to a reanalysis into phonological phrases in complex cases, the "readjustment rules" relating syntax to phonology make various other modifications in surface structure. It seems that in general these modifications involve elimination of structure, that is, deletion of nodes in representations such as (4) or of paired brackets in representations such as (5) (v. (4) e (5) na citação anterior). (...) These readjustment rules may somewhat modify the labeled bracketing of surface structure. They may also construct new feature matrices for certain strings of lexical and grammatical formatives. To take an obvious example, the verb sing will appear in the lexicon as a certain feature matrix, as will the verb mend. Using the letters of the alphabet as informal abbreviations for certain complexes of features, i. e., certain columns of a feature matrix, we can represent the syntactically generated surface structure underlying the forms sang and mended as $[v [v \text{ sing}] v \text{ past}] v$ and $[v [v \text{ mend}] v \text{ past}] v$, respectively, where past is a formative with an abstract feature structure introduced by syntactic rules. The readjustment rules would replace past by d, as a general rule; but, in the case of sang, would delete the item past with the associated labeled brackets, and would add to the i of sing a feature specification indicating that it is subject to a later phonological rule which, among other things, happens to convert i to æ. Designating this new column as *, the readjustment rules would therefore give the forms $[v s^*ng] v$ and $[v \text{ mend}] v$, respectively. We shall refer to this representation — and in general to the representation given by the application of all readjustment rules — as the "phonological representation". (Chomsky e Halle, 1968: 9-11).

Explica-se assim a maneira pela qual as regras do componente fonológico se aplicam a formas abstratas, ou seja, às representações léxicais e fonológicas. São exatamente estas as "formas básicas" a que já fizemos referência anteriormente.⁹ Quando frisamos não ser o seu estabelecimento de forma alguma

arbitrário, queremos dizer que tais formas são, em última análise, o output sintático. Isto explica também o motivo por que a análise fonológica pressupõe uma análise sintática já feita.

Resta ainda observar que a análise a partir de formas básicas abstratas implica em uma maior simplificação na gramática, uma vez que desaparece, na maioria dos casos, o problema da alomorfia. Os chamados alomorfes das análises estruturais explicam-se então naturalmente como sendo o resultado da aplicação de regras, independentemente motivadas na gramática de língua, às representações subjacentes. Como algumas das regras podem ainda fazer parte do chamado conjunto de regras universais (v. 1.1.3.1 e nota 5, deste capítulo), não representando portanto nenhum custo para a gramática específica que se esteja considerando, é natural que uma análise com base nos princípios teóricos aqui apresentados resulte mais econômica, simples e geral.

Apresentadas as bases teóricas para a análise do componente fonológico de uma gramática, passaremos à discussão dos complexos de traços distintivos que constituem o padrão sonoro do português.

1.2 Matrizes de traços distintivos para os segmentos do português.

Nas matrizes apresentadas nesta seção procuramos indicar quais seriam as combinações de traços fonéticos no nível abstrato das representações subjacentes (i.e., no nível das representações lexicais e fonológicas) dos enunciados do portu-

guês. A reserva com que apresentamos tais matrizes é perfeitamente justificável. Em uma língua específica, a matriz dos segmentos opositivos transcende, teoricamente, os limites de mera representação gráfica. O falante nativo, por possuir interiorizada a gramática de sua língua, conhece todos os complexos de traços por ela utilizados como segmentos opositivos. Estes segmentos são escolhidos pela língua dentre as várias combinações que constituem a matriz universal de representações fonéticas possíveis, de existência teoricamente prevista. Uma matriz específica reflete, portanto, o conhecimento por parte do indivíduo, dos segmentos opositivos de sua língua. Este conhecimento, que faz parte de sua competência linguística, reflete-se nos atos de fala.

Observe-se que, se a teoria universal da gramática já prevê todas as combinações de traços possíveis, o que faz uma gramática específica é escolher, dentre estas combinações, as que irão constituir a sua matriz. Esta escolha não é feita arbitrariamente, mas segundo princípios determinados que têm sido objeto de estudo nos últimos anos.¹⁰

Os diversos traços podem ser, em uma língua, distintivos ou redundantes, conforme sejam ou não utilizados para distinguir os segmentos entre si. É exatamente aqui que o linguísta enfrenta um problema de análise. Como saber a que traços atribui a língua que está analisando um caráter distintivo ou redundante?

A gramática de uma língua começa a existir juntamente com a língua e uma das grandes tarefas que se propõe a linguísti-

ca moderna é descobrir, descrever e sistematizar as gramáticas específicas. O mesmo é válido, em um âmbito mais restrito, com relação aos segmentos opositivos. As combinações de traços distintivos e redundantes em uma língua devem ser descobertas e não propostas. O trabalho do lingüista é, portanto, a descoberta de algo real, ainda que em um nível abstrato, e a matriz por ele apresentada para dar conta de todas as possibilidades de oposição no nível subjacente deve ser o reflexo da competência lingüística do falante nativo da língua. Dela devem fazer parte apenas os segmentos sentidos como subjacentes, excluindo-se todos aqueles cuja ocorrência, no nível da realidade, pode ser prevista por regras.

Todas estas observações a respeito da apresentação das matrizes se fazem aqui necessárias porque queremos deixar bem claro que é muito difícil chegar a uma descoberta desta espécie a partir da análise de uma pequena parte do componente fonológico, apenas. O analista terá maiores possibilidades de se aproximar da verdade subjacente, quanto maior for o corpus de que se utilizar para a análise. É claro que a situação ideal seria trabalhar, então, a partir de todos os dados da língua, tarefa que exigiria prazo muito mais longo do que o tempo de que dispusemos para o presente trabalho. Por serem o resultado do estudo de um corpus restrito, as matrizes que apresentamos para as consoantes e vogais do português têm caráter provisório, podendo sujeitar-se eventualmente a modificações que se mostrem necessárias em decorrência da consideração de novos problemas.

A especificação, nas matrizes que se seguem, é feita em termos estritamente binários, i.e., cada segmento, ao ser caracterizado com relação a determinado traço, recebe o valor + (se possui o traço) ou o valor — (se possui o traço oposto). Isto quer dizer que não se reconhecem valores intermediários ($\frac{+}{-}$) ou escalas graduadas de valores (traços n-ários) nas representações fonológicas.¹¹

1.2.1 Matriz das obstruintes e ressonantes

(1)

	p	t	k	b	d	g	f	s	š	v	z	ž	m	n	l	r
sonorante ¹²	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
consonantal ¹³	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	+	+	+	+
nasal	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	+	+	-	-
contínuo ¹⁴	-	-	-	-	-	-	+	+	+	+	+	+	(-)	(-)	+	-
coronal	-	+	-	-	+	-	-	+	+	-	+	+	-	+	(+)	(+)
anterior	+	(+)	-	+	(+)	-	(+)	+	-	(+)	+	-	(+)	(+)	(+)	(+)
tenso ¹⁵	+	+	+	-	-	-	+	+	+	-	-	-	(-)	(-)	(-)	(-)

Alguns dos segmentos acusam traços redundantes, que são indicados entre parênteses em (1).¹⁶ Estes traços são dispensáveis para a distinção entre os segmentos que compõem a matriz apresentada. Cada um dos segmentos opositivos encontra-se já perfeitamente distinguido dos demais pelo conjunto de

especificações a ele atribuído. Assim é que, no português, [g] abrevia o único segmento especificado como:

- | |
|-------------|
| - sonorante |
| - contínuo |
| - coronal |
| - anterior |
| - tenso |

A informação de que este segmento é também consonantal e não nasal é redundante no nível abstrato das representações fonológicas. Pode, no entanto, se mostrar necessária no nível do enunciado, quando se procura dar conta de todas as manifestações fonéticas ocorrentes na língua.

Certas observações se fazem aqui necessárias a respeito de alguns segmentos que, apesar de incluídos nos paradigmas estruturais dos fonemas do português, foram por nós excluídos da matriz de segmentos subjacentes da língua ou mantidos com alguma reserva:

(a) - Não consideramos segmentos subjacentes a nasal e lateral palatais. Preferimos interpretar a ocorrência, na fala, destes dois sons, como sendo resultante da aplicação de regras fonológicas a [n] e [l], respectivamente, em certos ambientes fonológicos. As regras em questão deverão ter seu âmbito de aplicação limitado a determinado estrato lexical, apenas, conforme se verá no capítulo 2;

(b) - Parece-nos também previsível a ocorrência do segmento que, no nível fonético, opõe-se a [r]. No diale-

to em que nos baseamos para a consideração de certos detalhes fonéticos, tal segmento se manifesta como uma fricativa glotal ou velar, dependendo do ambiente. Em outros dialetos manifesta-se como vibrante alveolar ou vibrante uvular;

(c) - $|\overset{v}{s}|$ e $|\overset{v}{z}|$ foram mantidos em (1) mas sabemos ser discutível o seu status de segmentos subjacentes. No entanto, somente a consideração de um corpus bem maior, abrangendo casos de derivação vocabular, poderia talvez confirmar a hipótese de que tais segmentos são fornecidos por regras, aplicáveis, como no caso de (a) e (b), a itens categorizados no léxico como especiais. Como o corpus com o qual trabalhamos é bastante restrito, não voltaremos a considerar o problema neste trabalho.

Com base nas especificações de (1) formularemos, no capítulo 3, as regras fonológicas que envolvem os segmentos incluídos nessa matriz.

1.2.2 Matriz das vogais

Segue-se a matriz das vogais do português:¹⁷

(2)

	i	e	ɛ	a	ɔ	o	u
sonorante	+	+	+	+	+	+	+
consonantal	-	-	-	-	-	-	-
posterior	-	-	-	+	+	+	+
alto	+	-	-	-	-	-	+
baixo	(-)	-	+	+	+	-	(-)
arredondado	(-)	(-)	(-)	-	+	(-)	(-)

Não reconhecemos motivos para postular a existência de glides subjacentes no português, pelo menos com base nos dados considerados. Explica-se assim a ausência de tais segmentos em (1) e (2) e também a possibilidade de caracterizar as obstruintes, nasais e vogais com apenas dois traços ("sonorante" e "consonantal") conforme fizemos. Se vier a ser provado que há necessidade de reconhecer glides subjacentes na língua, mais um traço ("vocálico") terá que ser considerado para dar conta das distinções (v. nota 13, deste capítulo). Partimos do princípio, na análise que apresentamos, de que as semivogais [y] e [w] resultam da aplicação de regras a [i] e [u], respectivamente, em circunstâncias a serem oportunamente indicadas.

1.3. Conclusão

Procuramos, neste capítulo, apresentar os princípios teóricos em que se fundamenta a análise comparativa do componente fonológico do português e do espanhol por nós apresentada no presente trabalho. As regras através das quais tentaremos caracterizar no capítulo 3 o mecanismo flexional de algumas das formas verbais do português bem como certos processos fonológicos de âmbito ^{mais} geral na língua, baseiam-se nas matrizes aqui introduzidas em (1) e (2). Deve ficar desde já esclarecido, no entanto, que podem se operar modificações, às vezes radicais, em uma matriz, por meio de regras que invertem o valor de certos traços. Isto pode ser necessário para que se expliquem certas manifestações fonéticas.¹⁸ No nível das represen

tações subjacentes, porém, os elementos caracterizam-se por possuir sempre as mesmas marcas fonêmicas abstratas, assegurando-se assim o funcionamento do sistema de oposições da língua em questão.

Uma observação final deve ainda ser feita com relação ao acento de intensidade. Considerado fonema supra-segmental nas análises estruturais da língua por distinguir itens segmentalmente idênticos (cf. a exemplificação clássica sábia (adj.) : sabia (verbo) : sabiá (subst.), é aqui interpretado como um fenômeno supra-segmental condicionado, entre outros fatores, pela natureza de certos segmentos (vogais) sobre os quais ele pode ou não incidir. Isto significa que é também previsível, no português, e decorrente da aplicação de regras. Voltaremos a abordar o problema do acento em outras seções do trabalho, embora não de maneira exaustiva, já que para isso seria necessário examinar um corpus bem maior e que incluísse também itens representativos de outras classes gramaticais. Trataremos aqui, portanto, do problema da acentuação no português na medida apenas em que for relevante para a consideração dos dados que nos propusemos analisar, ficando um estudo mais completo para um trabalho também mais inclusivo.

No capítulo seguinte trataremos da necessidade de identificar estratos lexicais no português, estratos estes que, como veremos a seguir, virão a constituir âmbito de aplicação para as diversas regras fonológicas.

CAPÍTULO 2

OS ESTRATOS LEXICAIS E O ÂMBITO DE APLICAÇÃO DAS REGRAS FONOLÓGICAS

2.1 Considerações preliminares

Observem-se os seguintes grupos de dados do português:

(1)	I	II
A.	lunar	lua
	monetário	moeda
	arenoso	areia
	venoso	veia
	donativo	doação
B.	voluntarioso	vontade
	colorido	cor
	solitário	só
	salutar	saúde
	filamento	fio
C.	manual	mão
	canil	cão
	saponáceo	sabão
	granulado	grão
	racional	razão
D.	unitário	unidade
	fatal	fado
	rotação	roda
	vital	vida
	mutismo	mudo
E.	sacro	sagrado
	vacante	vago
	local	lugar
	mendicância	mendigo

E.

I

articulista

II

artigo

F.

apicultura

abelha

saponáceo

sabão

capital

cabeça

leporino

lebre

capilar

cabelo

G.

plano

chão

pleno

cheio

clave

chave

clamar

chamar

flama

chama

H.

capital

cabeça

apicultura

abelha

hospital

hóspede

digital

dedo

signo

senha

I.

factual

feito

lácteo

leite

sectário

seita

expectorar

peito

noctívago

noite

J.

bucal

boca

lucro

logro

cúpreo

cobre

super

sobre

sumário

soma

com base no exame dos dados, pode-se dizer que:

(a) - há um denominador comum semântico entre cada item da coluna I e seu correspondente da coluna II, nos diversos grupos. São diferentes, no entanto, as configurações fonológicas das raízes em questão. As principais diferenças formais são:¹

(2)	I	II
A.	n	∅
B.	l	∅
C.	n	VE [ãw]
D.	t	d
E.	c [k]	g
F.	p	b
G.	pl cl fl	ch [s]
H.	ac [ak] ec [ek] oc [ok]	ei [ey] oi [oy]
I.	i	e
J.	u	o ;

(b) - se, tomando por base a semelhança semântica, conclui-se haver uma só raiz para cada par de itens assim relacionados, terão que ser admitidas, na gramática do português, regras fonológicas aplicáveis apenas a formativos como os da coluna II. Estas regras serão, em última análise, responsáveis pelas diferenças formais observadas acima;

(c) - se nem todos os formativos estão sujeitos à aplicação de todas as regras, a gramática da língua deverá indicar, de algum modo, o âmbito de aplicação das regras fonológicas;

(d) - é possível pensar em uma estratificação do vocabulário, identificando-se categorias de formativos às quais se aplicam ou não os conjuntos de regras.

Na seção seguinte consideraremos o problema da estratificação lexical no português.

2.2 O reconhecimento de estratos lexicais no português

Vimos já como vários formativos semanticamente relacionados no português, apresentam um comportamento diferente com relação a certos processos fonológicos (queda de segmento nasal ou lateral entre vogais, lenização de oclusivas também em ambiente intervocálico, mudança na qualidade de certas vogais, etc.).² Enquanto as formas da coluna II são o resultado da operação destes processos, as formas da coluna I não são afetadas pelas regras. É, aliás, o próprio fato de coexistirem na língua pares de formas como as apresentadas, que sugere a existência de regras operantes em certas categorias de formativos, apenas. Com efeito, não seria necessário postular uma regra sincrônica que sonorizasse as oclusivas intervocálicas, não fosse o fato da gramática ter que dar conta da existência de pares como unitário e unidade, sacro e sagrado, apicultura e abelha, no português atual. A regra é sincronicamente motivada, portanto, aplicando-se às formas da coluna I, apenas.

É importante ressaltar que muitos outros dados podem ser acrescentados às duas listas, constituindo-se inclusive novos grupos pela identificação de outros processos, sem que as formas de I e II percam sua característica básica com relação a determinado conjunto de regras (i.e., regras aplicáveis a II,

mas não a I). Exemplos de novos grupos de formas pertencentes a I e II, seriam:

(3)	I	II
L.	pedestal	pé
	sede	sé
	crédito	crer
	vidente	ver
	sorridente	sorrir
	(d)	(Ø)
M.	legível	ler
	digital	dedo
	frígido	frio
	rugido	ruído
	sigilo	selo
	(g [ž])	(Ø)

Observe-se que a relação é sempre a mesma. A queda dos segmentos indicados verifica-se nas formas da coluna II, conforme se deduz do confronto com as formas de I.

I e II representam, pois, por seu comportamento distinto, amostra de dois estratos lexicais existentes no português. Os formativos que constituem um dos dois estratos devem vir marcados no léxico para que se possam aplicar convenientemente as regras do componente fonológico da gramática.³ Para fins de economia, receberão marca apenas os formativos que constituem o menor grupo, considerando-se I e II com relação a todo o léxico do português. É empiricamente verificável a afirmação de que constituem maioria na língua os formativos do tipo I, não

sujeitos à aplicação de certas regras.⁴ Serão marcados, portanto, os formativos do tipo II, como sujeitos às regras.

Trataremos a seguir da natureza das marcas a serem atribuídas aos formativos e da maneira como devem ser interpretadas com relação aos segmentos de cada matriz lexical.

2.3 As propriedades diacríticas

Os formativos podem vir marcados, no léxico, por certas categorias sintáticas e semânticas (e. g. [nome], [animado], [humano], etc.). Explicam-se desta forma as categorizações que se estabelecem nas línguas e em função das quais se faz, muitas vezes, a aplicação de regras gramaticais.

Ao lado destas categorias, os formativos podem também ser marcados com relação a propriedades diacríticas, necessárias para o estudo das chamadas "exceções". Considerem-se, a esse respeito, alguns trechos de Chomsky e Halle (1968):

Many grammatical rules apply only to certain lexical items. For example, in a language with a rich inflectional system, such as Latin or Russian, it is necessary to divide all noun stems into several declensional classes to account for the phonetic realization of the gender, number, and case features. These classes may play no other role in the grammar; in particular, they generally have no syntactic function. We shall represent this rather peripheral classification with special "diacritic features" in lexical entries. (...)

In the phonology proper, we also find quite commonly that rules apply in a selective fashion and thus impose an idiosyncratic classification on the lexicon. Often there is a historical explanation for this idiosyncratic behavior, but this is obviously irrelevant as far as the linguistic competence of the native speaker is concerned. What the speaker knows is, simply, that a given item or set of items is treated differently

from others by the phonological component of the grammar (373). (Grifo nosso)

Each phonological rule of the language applies to certain formatives and, in general, not to others, the domain of its application being determined by the feature composition of the phonological matrices. If a certain rule does not apply to a certain formative, this fact must somehow be indicated in the feature composition of the formative at the stage of derivation at which the rule is applicable. It is quite obvious that many of the phonological rules of the language will have certain exceptions which, from the point of view of the synchronic description, will be quite arbitrary. This is no more surprising than the fact that there exist strong verbs or irregular plurals. Phonology, being essentially a finite system, can tolerate some lack of regularity (exceptions can be memorized); being a highly intricate system, resulting (very strikingly, in a language like English) from diverse and interwoven historical process, it is to be expected that a margin of irregularity will persist in almost every aspect of the phonological description. Clearly, we must design our linguistic theory in such a way that the existence of exceptions does not prevent the systematic formulation of those regularities that remain. Furthermore, we must provide means for expressing those regularities that hold within the class of exceptions, however limited they may be. Finally, an overriding consideration is that the evaluation measure must be designed in such a way that the wider and more varied the class of exceptions to a rule, the less highly valued is the grammar (172).

The natural way to reflect such exceptional behavior in the grammar is to associate with such lexical items diacritic features referring to particular rules (374).

Após justificar a inclusão das propriedades diacríticas nas gramáticas das línguas, tratam os autores da maneira como são introduzidas e interpretadas estas propriedades dos formativos. Abordaremos aqui somente os aspectos que nos parecem essenciais para a compreensão da análise que apresentamos no capítulo 3.

As propriedades diacríticas estabelecem também categorizações no léxico, pois são atribuídas diretamente aos formativos, que constituem o seu domínio.⁵ Constituem-se assim subclasses morfológicas definíveis em termos de processos fonológicos características. Por convenção, todas as categorias lexicais a que pertence um formativo tornam-se propriedades distintivas de cada um dos seus segmentos. Ao se aplicar esta convenção verifica-se que as marcas recebidas pelos segmentos com relação às propriedades sintáticas e semânticas não têm, via de regra, nenhum efeito fonético. Já as propriedades diacríticas podem ser diretamente responsáveis por modificações no output fonético, na medida em que os segmentos dos diversos formativos que constituem o léxico da língua sejam ou não marcados com relação a estas propriedades.

Empregaremos, no presente trabalho, a propriedade diacrítica [E] para caracterizar os formativos do português como pertencentes a um ou outro dos estratos lexicais já identificados. [E] significa "especial" e utilizamos este termo, tal como Harris na análise do Espanhol (Harris, 1969 : 127-30) , porque nos parece mais livre de conotações (v. 2.4).

Receberão marca no léxico, conforme já vimos (2.2), os formativos do tipo II. Serão, portanto, [mE], enquanto que os formativos do tipo I serão não marcados, i.e., [uE]. Uma convenção universal de interpretação desta propriedade (ou equivalente) dirá que:

$$\begin{array}{lcl} [mE] & \longrightarrow & [+E] \\ [uE] & \longrightarrow & [-E] \end{array}^6$$

Por coincidir o emprego que fazemos de $[E]$ com o que faz Harris no trabalho citado, transcrevemos aqui um trecho em que o autor explicita a maneira de interpretar as marcas $[+E]$ e $[-E]$:⁷

(a) All formatives are subcategorized with respect to the morphological feature $[S]$, mnemonic for "special": all formatives are either $[+S]$ or $[-S]$.

(b) Formatives subcategorized as $[+S]$ are subject to all phonological rules that $[-S]$ formatives are subject to. Formatives subcategorized as $[+S]$ are also subject to an additional designated subset of rules that $[-S]$ formatives are not subject to.

(c) Rules that apply to $[+S]$ formatives only are so designated by the inclusion of $[+S]$ in the structural description. That is, such rules are of the form $[A, +S] \longrightarrow B/X \text{ --- } Y$ or, equivalently, $A \longrightarrow B/X \text{ --- } [+S] Y$, where A, B, X, Y are interpreted in the usual way. These rules are not all contiguously ordered; They are interspersed among other rules.

(d) Items in the lexicon are either marked or unmarked for $[S]$. By convention, $[mS] \longrightarrow [+S]$, $[S] \longrightarrow [-S]$ (a glance through a dictionary is sufficient to show that the vast bulk of the lexicon is $[-S]$, that is, unmarked. Heuristically, a lexical formative is assumed to be unmarked until one is forced to the conclusion that it is marked.)

(e) Inflectional formatives (complexes of features for person, number, mood, aspect, etc.) are automatically assigned the feature $[+S]$. The features $[+S]$, $[-S]$ are features of formatives, not words: $[+S]$ affixes may be attached to $[-S]$ stems.

Ressaltamos, do trecho acima, dois aspectos importantes: o fato dos formativos marcados com $[+E]$ estarem sujeitos também a todas as regras que se aplicam aos formativos $[-E]$ e o

fato de formativos flexionais (automaticamente [+E]) poderem se combinar também com formativos [-E] (radicais, no caso). A partir da consideração do item (a) acima pode-se concluir, portanto, que há uma série de regras gerais aplicáveis a todos os formativos da língua e um conjunto de regras especiais que se aplicam apenas à categoria [+E] de formativos. O item (e) torna claro o motivo por que a propriedade [E] aplica-se a formativos, e não a palavras. Palavras podem apresentar mais de um formativo podendo combinar-se, portanto, em uma só palavra, um formativo [-E] com um formativo [+E]. Este é o caso, por exemplo, dos radicais [-E] que se combinam, no português, com os formativos flexionais (que são, automaticamente, [+E]).⁸

A consideração de todos os dados da língua levará, fatalmente, à conclusão de que o reconhecimento de apenas dois estratos lexicais no português é insuficiente para a formulação adequada de todas as regras fonológicas. Grupos de exceções aparentes se explicam automaticamente, uma vez admitidos outros substratos lexicais (e, conseqüentemente, conjuntos de regras aplicáveis a eles apenas).⁹ Neste trabalho não tentaremos, no entanto, apresentar tais subcategorizações, uma vez que não será possível trabalhar com todos os dados da língua e qualquer tentativa de uma subcategorização exaustiva do léxico, a partir de um corpus restrito como o nosso, seria apriorística e provavelmente inexata. O reconhecimento dos dois estratos básicos I e II, respectivamente caracterizados como [-E] e [+E], é suficiente para o estudo que aqui nos pro

pomos fazer.

2.4 Sincronia e diacronia

Achamos perfeitamente justificável dedicar uma seção deste capítulo à consideração das relações existentes entre um estudo sincrônico e um estudo diacrônico da língua. Com efeito, após a apresentação de todos os argumentos para o estabelecimento dos dois grandes estratos lexicais do português, difícil seria não associar o estrato constituído de formativos especificados como [+E], como o estrato que sabemos ser, historicamente, constituído dos itens lexicais que evoluíram normalmente do latim vulgar para o português. Da mesma forma, somos imediatamente forçados a reconhecer que pertencem ao outro estrato (formativos [-E]) os itens que constituam empréstimos eruditos do latim, por exemplo, introduzidos posteriormente na língua (cf. plano ([-E]) e chão ([+E])).

Se é possível estabelecer categorizações, com respeito a dados de um estado atual de língua, que confirmam fatos historicamente comprovados, bem como regras sincrônicas que são também comprovadamente diacrônicas, impõe-se a pergunta: até que ponto é válida a utilização de informações históricas para a análise de fatos sincrônicos? O problema nos parece simples. Não se deve confundir a distinção feita sincronicamente entre [+E] e [-E] com a distinção historicamente atestada entre itens "vulgares" e "eruditos", embora a primeira possa ser, em muitos casos, um reflexo da última. Nos grupos de exemplos apresentados em (1) existe, de fato, uma coincidência absoluta.

Ela não é, no entanto, necessariamente válida para todos os formativos da língua. Um item lexical, por exemplo, que é sincronicamente [-E] por não se submeter ao conjunto de regras aplicáveis apenas aos itens [+E], pode ser "vulgar" do ponto-de-vista diacrônico.

Pode acontecer também que uma informação histórica seja totalmente irrelevante para a consideração de determinado problema sincrônico, conforme se depreende do seguinte trecho de Schane (1968):

It is important that the learned-nonlearned distinction within a synchronic description not be confused with the historical events of which this distinction is to a large extent a reflexion. Although the historical events may provide an explanation for the alternations attested in the contemporary language, the synchronic description itself does not depend on knowledge of such facts. The grammatical category of gender illustrates this point very well. It is of historical interest to know that French champ 'field' is masculine and table 'table' is feminine because the Latin words campus and tabula from which these developed were masculine and feminine respectively. However, this piece of historical information is totally irrelevant for establishing gender classification for contemporary French since the gender distinction in French can be internally motivated morphologically and syntactically. (...) Since these divisions are established on internal grounds—morphological and phonological evidence within modern French—they may not necessarily correlate with historical developments within the language. That the effects of such division often do coincide with historical facts should not be surprising, however. (grifo nosso). Within the description of French presented here, the terms "learned" and "non-learned", although descriptive, are nonetheless arbitrary designations of classes (cf. Type 1 and type 2 used earlier), and ought to be understood in the same way as other traditionally used terms, such as masculine, first conjugation, inflectional, and so on. (27-8).

O fato é que tanto um estudo diacrônico como um estudo sincrônico devem, por definição, ser auto-suficientes. Não é lícito, por exemplo, que o lingüista, para fins de simplificação de uma análise que se propõe ser sincrônica, estabeleça uma estratificação lexical baseada em informações diacrônicas. Em um estudo sincrônico, tal estratificação só se justifica e admite se sincronicamente motivada.¹⁰ A estratificação por nós proposta para o português é perfeitamente motivada do ponto-de-vista sincrônico, uma vez que coexistem, no estado da língua com que trabalhamos, as formas I e II (v. 2.1). O simples exame dos dados sincrônicos leva, portanto, à categorização proposta.

Não se pode negar, no entanto, que o conhecimento dos fatos históricos pode ajudar a analista a vislumbrar hipóteses para a explicação de fatos sincrônicos. Outra coisa não seria de se esperar, aliás. O essencial é, frisamos mais uma vez, mostrar que as hipóteses são justificáveis a partir da consideração de dados sincrônicos, apenas. Em trabalho sobre a fonologia e morfologia do francês, Schane manifesta-se precisamente sobre este problema:

(...) From a historical point of view an /o/ to /u/ conversion is completely justified. However, we believe that synchronic rules must be internally motivated and that their sole justification can not be the corresponding diachronic rules, although, to be sure, the latter certainly help to corroborate the validity of the former. In working out an analysis, it is unquestionably an advantage to know the historical development of the language just as it is advantageous to have access to the standard orthography (for a written language) and to have in-

formation on related dialects and languages. To ignore completely historical documents in a language extensively documented as French would be foolish indeed. Knowing the history of the language allows one to formulate hypotheses concerning the nature of underlying forms and the types of phonological changes still operative in the contemporary language. The confirmation of these hypotheses must then be demonstrated uniquely within the synchronic description. Such internal synchronic justification is indispensable if linguistic descriptions are to have psycholinguistic import and are to characterize what speakers know about their language. (grifo nosso). (Schane, 1968 : 140-1).

A respeito da propriedade diacrítica ([E]), que empregamos para categorizar os formativos no léxico, dissemos, alguns parágrafos atrás, que a preferência pelo termo "especial" se deveu ao fato de nos parecer ele mais livre de conotações. Poderíamos, evidentemente, ter empregado "vulgar" (por [+E]) e "erudito" (por [-E]), ou algo nesta linha.¹¹ Se rejeitamos tais termos, foi porque eles estão tradicionalmente relacionados com processos históricos, não sendo, portanto, adequados a uma análise que procura justamente demonstrar a independência de um estudo sincrônico. Além disso, optando pela terminologia "especial", de certa forma mais neutra, tornamos menos restrita a categoria dos formativos que não se submetem a determinado conjunto de regras. Assim, podem ser considerados [-E] não apenas os formativos historicamente "eruditos", mas também empréstimos de outras línguas além do Latim, bem como os demais itens que, por qualquer motivo, não se comportam como os formativos que constituem a categoria [+E], muito mais limitada (formativos historicamente "vulgares"). Es-

ta divisão primária da língua em dois estratos permite uma grande simplificação na análise. Uma vez isolados os formativos "especiais" pode-se trabalhar simultaneamente na identificação das regras gerais (aplicáveis a todo o léxico) e do subconjunto de regras aplicáveis aos formativos desse estrato, apenas. Uma etapa seguinte seria subcategorizar o estrato $[-E]$, estabelecendo novos substratos lexicais de acordo com o comportamento dos formativos $[-E]$ com relação a outros prováveis subconjuntos de regras.

Uma primeira abordagem do componente fonológico das diversas línguas resulta, naturalmente, na consideração das regras gerais e das que incidem sobre os formativos $[+E]$. Isto é perfeitamente explicável, partindo-se do princípio de que oferece maior interesse o estudo do estoque de itens que constituem, de certa forma, o verdadeiro vocabulário "nativo" de qualquer língua considerada. É, aliás, pelo mesmo motivo, que os estudos diacrônicos tradicionais costumavam se fixar também nesse estrato, enfatizando sempre o processo de evolução dos formativos, a partir de proto-línguas atestadas ou reconstruídas. Observe-se ainda que, devido à ênfase dada ao estudo dos formativos $[E]$, obscureceu-se o fato de que esta categoria é minoritária em muitas línguas (português, por exemplo), com relação ao número de itens de que é constituída.

Conclui-se, portanto, que análises sincrônicas e diacrônicas podem e devem ser feitas isoladamente. A comparação entre os resultados de tais estudos é, no entanto, perfeita-

mente admissível, sendo de se esperar inclusive que fatos históricos atestados confirmem, até certo ponto, a análise de um estado sincrônico de língua (v. nota 10 deste capítulo).

2.5 Conclusão

Pode-se concluir o presente capítulo com um resumo do que ficou estabelecido a respeito da estratificação lexical no português.

Vimos que a estratificação do léxico é indispensável para uma análise adequada do componente fonológico da língua e dividimos os formativos do português em dois grandes estratos, com base no âmbito de aplicação de um determinado conjunto de regras fonológicas. Consideramos a seguir as propriedades diacríticas e algumas de suas características formais, utilizando a propriedade "especial" para marcar os dois estratos como $[-E]$ e $[+E]$. No capítulo seguinte veremos que as regras fonológicas com a especificação $[+E]$ na descrição estrutural são aplicáveis apenas aos formativos $[+E]$.

Tratamos ainda da espécie de relação existente entre estudos sincrônicos e diacrônicos. Mostramos serem perfeitamente autônomos, devendo haver sempre uma justificação sincrônica para fatos sincrônicos, e diacrônica para fatos diacrônicos. Ressaltamos a validade de posterior comparação entre uma análise sincrônica e uma análise diacrônica de determinada língua. Consideramos, por fim, a possibilidade de confirmação de hipóteses sobre processos sincrônicos através do confronto com processos historicamente comprovados.

Os princípios teóricos comentados nestes capítulos introdutórios servirão de base para a análise comparativa entre formas verbais do português e do espanhol que passaremos a expor no capítulo seguinte.

CAPÍTULO 3

AS FORMAS DO PRESENTE DO INDICATIVO NO ESPANHOL
E NO PORTUGUÊS: EVIDÊNCIAS PARA IDENTIDADE DE
REPRESENTAÇÕES BÁSICAS

3.1 Considerações preliminares

Nos capítulos precedentes tivemos oportunidade de mencionar por diversas vezes a obra de Harris (1969) Spanish Phonology. Com efeito é de se esperar que coincidam, até certo ponto, as observações que se podem fazer a respeito do português e do espanhol, dada a estreita relação genética entre as duas línguas. Historicamente é explicável a semelhança observada. O que pretendemos examinar neste capítulo, no entanto, é a possibilidade de considerar válido tal relacionamento no plano sincrônico, sem que seja necessário buscar evidências de ordem histórica. Pretendemos, em suma, argumentar a favor da hipótese de que se pode partir de formas básicas idênticas, na maioria dos casos, para a análise do componente fonológico das duas línguas. Tal hipótese implica na consideração das diferenças superficiais como sendo o resultado da aplicação de maior número de regras a uma das línguas, da ordenação diferente de um mesmo conjunto de regras, da supressão de regras, etc. Estes seriam, em última análise, os fatores responsáveis pela diferenciação superficial entre o português e o espanhol. A esse respeito é interessante notar que são justamente as regras mais profundas, consideradas básicas, que se aplicam em condições quase sempre idênticas às duas línguas.

Observa-se logo que, ao operar com formas básicas e uma série de regras, obtêm-se em um estágio bastante próximo do nível fonológico subjacente as manifestações fonéticas para o espanhol, devendo-se continuar com a aplicação de outras regras para chegar às representações superficiais do português.

É justamente a identificação e formulação adequada de tais regras que constitui um dos problemas a serem enfrentados pelo analista que se propõe analisar o componente fonológico desta língua. Quais serão, por exemplo, as regras mais naturais para explicar os chamados ditongos nasais? Não é difícil imaginar um conjunto de regras que, de alguma forma, dêem conta do problema. O caso é que, na maioria das vezes, as regras parecem ad hoc, i.e., carecem de motivação independente na língua, sendo formuladas com vistas à resolução de problemas específicos, apenas.

Não é nosso propósito apresentar aqui soluções definitivas para cada problema surgido da consideração de certas alternâncias regulares identificadas a partir do exame das formas verbais em questão. Propomo-nos, antes de tudo, a apresentar de modo claro esses problemas, que são relativamente numerosos, índice talvez de uma gramática um pouco mais complexa no caso do português em confronto com o espanhol. É importante notar que vários estudiosos da fonologia gerativa têm sido forçados a observar que alguns exemplos do português poderiam levar talvez à reformulação de certos princípios teóricos da disciplina. Com efeito, o trabalho com línguas específicas tem contribuído para a identificação de aspectos teóricos que merecem re-exame e possível reformulação.¹ Não poderia, aliás, ser diferente. Uma teoria não é concebida em sua forma ideal, livre de imperfeições. Estas imperfeições só podem ser identificadas e corrigidas após tentativas de aplicação da teoria. O presente estudo constitui exatamente uma tentativa de aplicação da teoria fonológica gerativa a uma análise comparativa

de algumas formas verbais do espanhol e do português. Restringimo-nos às formas do presente do indicativo das três conjugações regulares, principalmente por não termos tido tempo suficiente para considerar as demais formas verbais. Achamos, no entanto, que a análise destas formas apenas permite já que se delineiem os rumos a serem tomados em um trabalho de âmbito mais geral. Teremos alcançado plenamente o nosso objetivo se conseguirmos apontar sistematicamente os problemas com os quais nos deparamos ao examinar o nosso corpus limitado. Esperamos que de uma futura consideração conjunta de problemas como os que serão discutidos nos parágrafos seguintes possam resultar sugestões valiosas para o desenvolvimento da teoria em geral (i.e., quando estes problemas forem estudados tomando-se por base um corpus mais amplo de formas de uma mesma língua ou de línguas diferentes).

Antes de passarmos propriamente à comparação das formas, apresentaremos na seção seguinte alguns aspectos característicos da classe verbal no português.

3.2 A classe verbal no português

O verbo talvez seja a classe gramatical que maior atenção tem merecido por parte dos estudiosos da língua portuguesa. Ao lado do grande número de trabalhos tradicionais alinham-se várias obras de cunho estruturalista, nas quais se procura sistematizar o padrão flexional do verbo português de acordo com os princípios teóricos desta escola linguística.²

Atualmente existe interesse em considerar a classe verbal do

ponto-de-vista da teoria transformacional, o que equivale a dizer que se procuram identificar as alternâncias regulares dentro desta classe e as regras que dão conta de tais regularidades. Buscam-se principalmente os processos que operam sobre formas consideradas básicas e que são responsáveis por certas manifestações fonéticas.³

Considerado o "vocabulário flexional por excelência" (Camara, 1970 : 94), o verbo apresenta flexões de número, pessoa, modo, tempo e aspecto. Essas noções gramaticais, sobre as quais falaremos resumidamente a seguir, exprimem-se, na grande maioria dos casos, através de formativos gramaticais acrescentados às bases verbais (formativos lexicais acrescidos da vogal característica da base). Podem também por vezes vir expressas por alternâncias fonológicas nas próprias bases verbais, como é o caso de certas noções de pessoas (cf. fiz : fez) e tempo (cf. pode : pôde).

O resultado das várias possibilidades de combinação destas noções gramaticais a cada uma das bases verbais vem a ser o grande número de formas verbais existentes na língua. Algumas destas formas serão objeto de estudo deste capítulo, quando abordaremos problemas como:

(a) - a maneira de caracterizar formativos gramaticais que exprimem as noções gramaticais de número, pessoa, modo, tempo e aspecto;⁴

(b) - a maneira como se combinam, nas representações fonológicas, estes formativos gramaticais com os formativos lexicais que, acrescidos da vogal característica, cons

tituem as bases verbais;

- (c) - a aplicação, a estas representações fonológicas, das regras do componente fonológico que explicam as manifestações fonéticas ocorrentes na cadeia da fala.

Ao tratar destes problemas com relação ao presente do indicativo tentaremos, conforme já observamos, estabelecer uma comparação com a análise do espanhol proposta por Harris. Em vários pontos haverá coincidência. Em outros, veremos que há necessidade de ir um pouco mais além no português ou mesmo de propor alguma maneira diferente de abordar determinado problema.

A análise de base transformacional que introduziremos nas seções seguintes tem como ponto de partida principalmente a análise estrutural proposta por Camara (1969, 1970) que apresenta já uma sistematização dos formativos e de suas possibilidades de combinação dentro do vocábulo verbal.

3.2.1 As noções gramaticais de número, pessoa, tempo, aspecto, modo e voz.

Trataremos brevemente a seguir das noções gramaticais de número, pessoa, tempo, aspecto, modo e voz. Achamos oportuno introduzir tais noções antes da apresentação da análise propriamente dita, já que pode eventualmente ser necessário fazer menção a algumas delas.

Deve ficar desde já esclarecido que a noção gramatical de voz, considerada pela NGB como uma das "flexões" verbais, expressa-se, na realidade, através de um processo sintático, no português. Não há, com efeito, nenhum formativo gramatical

na língua ao qual se possa atribuir tal noção.⁵

3.2.1.1 Número

O número expressa, na língua, as noções de singular e plural. Nos verbos, é índice da singularidade ou pluralidade da frase nominal, sujeito da frase verbal.

Note-se que em português uma frase nominal de conteúdo semântico plural e forma singular (e. g. uma FN cujo núcleo é um coletivo) faz-se acompanhar de um verbo na forma singular, porque é a forma, no caso, o fator determinante da concordância. O mesmo se observa com as FN's de conteúdo semântico singular e forma plural, que são seguidas de forma verbal no plural.⁶

O formativo que carrega a noção de número nas formas verbais portuguesas (sufixo número-pessoal - SNP) expressa também a noção de pessoa, conforme veremos a seguir.

3.2.1.2 Pessoa

A noção gramatical de pessoa faz referência, nos verbos, às características da FN sujeito como participante do processo de comunicação. Mencionam-se seis subdivisões possíveis de pessoa, conforme a frase nominal representante:

o falante ou P(essoa) 1, o falante e mais alguém ou P(essoa) 4, um ouvinte ou P(essoa) 2, mais de um ouvinte ou P(essoa) 5, um ser ou mais de um ser distintos do falante e do ouvinte, ou seja, respectivamente, P(essoa) 3 e P(essoa) 6. (Camara, 1970 : 94-5).

A noção de pessoa está intimamente relacionada à noção de número, já que a frase nominal sujeito da frase verbal pode

ser, conforme vimos (v. 3.2.1.1), singular ou plural. As duas noções, portanto, implicam em referências cruzadas, o que explica o fato de virem sempre expressas cumulativamente na forma verbal, através de um mesmo formativo gramatical.

No decorrer da análise, para facilidade de exposição, faremos sempre referência às fórmulas propostas por Camara (1969, 1970), ou seja: P_1 , P_2 , P_3 , P_4 , P_5 , P_6 .

3.2.1.3 Tempo⁷

A noção gramatical de tempo refere-se ao conteúdo temporal das formas verbais, i.e., à relação entre estas formas e o momento no tempo em que decorre a ação.

Há dois eixos temporais básicos: passado e não passado. No eixo não passado costumam-se distinguir as categorias de presente e futuro. O presente diz respeito à ação que se realiza a partir do momento da fala. O chamado tempo futuro possui, no português, certas conotações "modais" (v. 3.2.1.5): implica em um desejo ou certeza, do ponto-de-vista do falante, de que determinado evento se realize (do ponto-de-vista do presente) ou se tenha realizado (do ponto-de-vista do passado). A categoria de futuro dá margem, na língua, a um estudo bastante detalhado, do qual não nos ocuparemos no momento.⁸ Os eixos temporais passado e não passado, juntamente com os aspectos (v. 3.2.1.4), dão origem aos complexos de formas que se denominam "tempos" dentro do sistema flexional do verbo. Estes "tempos", por sua vez, organizam-se tradicionalmente em "modos" (v. 3.2.1.5).

3.2.1.4 Aspecto

A categoria de aspecto é, sem dúvida, sincronicamente operante no sistema verbal do português. No entanto, embora se reconheça o caráter funcional desta noção na gramática da língua, há bem poucos estudos dedicados ao assunto, até o presente. Isto se deve, em parte, ao fato de estar o aspecto intimamente relacionado ao tempo, nas formas verbais, o que torna muitas vezes difícil a exata identificação e sistematização destas duas noções. Na verdade, "Todas as formas do verbo são combinações de tempo e aspecto" (Rallides, 1971 : 9).

Distinguem-se, nas formas verbais portuguesas, dois aspectos principais : perfectivo e imperfectivo (i.e., ação conclusa ou não conclusa, do ponto-de-vista do falante). Estas duas categorias aspectuais determinam classes morfológicas nos verbos, identificáveis a partir do exame de certos verbos considerados "irregulares" (e. g. observando-se o radical de fazer nota-se que a vogal radical é [a] nas formas que se caracterizam por apresentarem aspecto imperfectivo, e [i] ou [e] quando o aspecto é perfectivo). Diferenças de aspecto podem, portanto, fornecer um critério formal para o agrupamento das formas verbais na língua.

3.2.1.5 Modo

A noção gramatical de modo é geralmente definida segundo um critério semântico, i.e., diz-se que o modo exprime a maneira como se realiza o fato verbal.

Identificam-se três modos: indicativo, subjuntivo e impe-

rativo. A "maneira como se realiza o fato verbal" seria, nestes três modos:

- (a) - indicativo: referência, de forma categórica; segura, a fatos reais; afirmação, ou negação de uma realidade;
- (b) - subjuntivo: expressão de hipótese, desejo, probabilidade, dúvida, incerteza;
- (c) - imperativo: expressão de ordem, vontade, exortação.

Deve-se observar, contudo, que nem sempre o conteúdo modal de uma sentença é expresso formalmente através do verbo. A noção de modo é, muitas vezes, função da própria estrutura da frase.

3.2.1.6 Voz

A categoria de voz refere-se ao papel do sujeito com relação ao processo verbal. Em linhas gerais, pode-se dizer que há três vozes verbais:

- (a) - voz ativa: quando o sujeito é o agente da ação expressa pela forma verbal;
- (b) - voz passiva: quando o sujeito é o paciente da ação verbal;
- (c) - voz reflexiva: quando o sujeito é ao mesmo tempo agente e paciente, tomando-se como base o processo verbal.

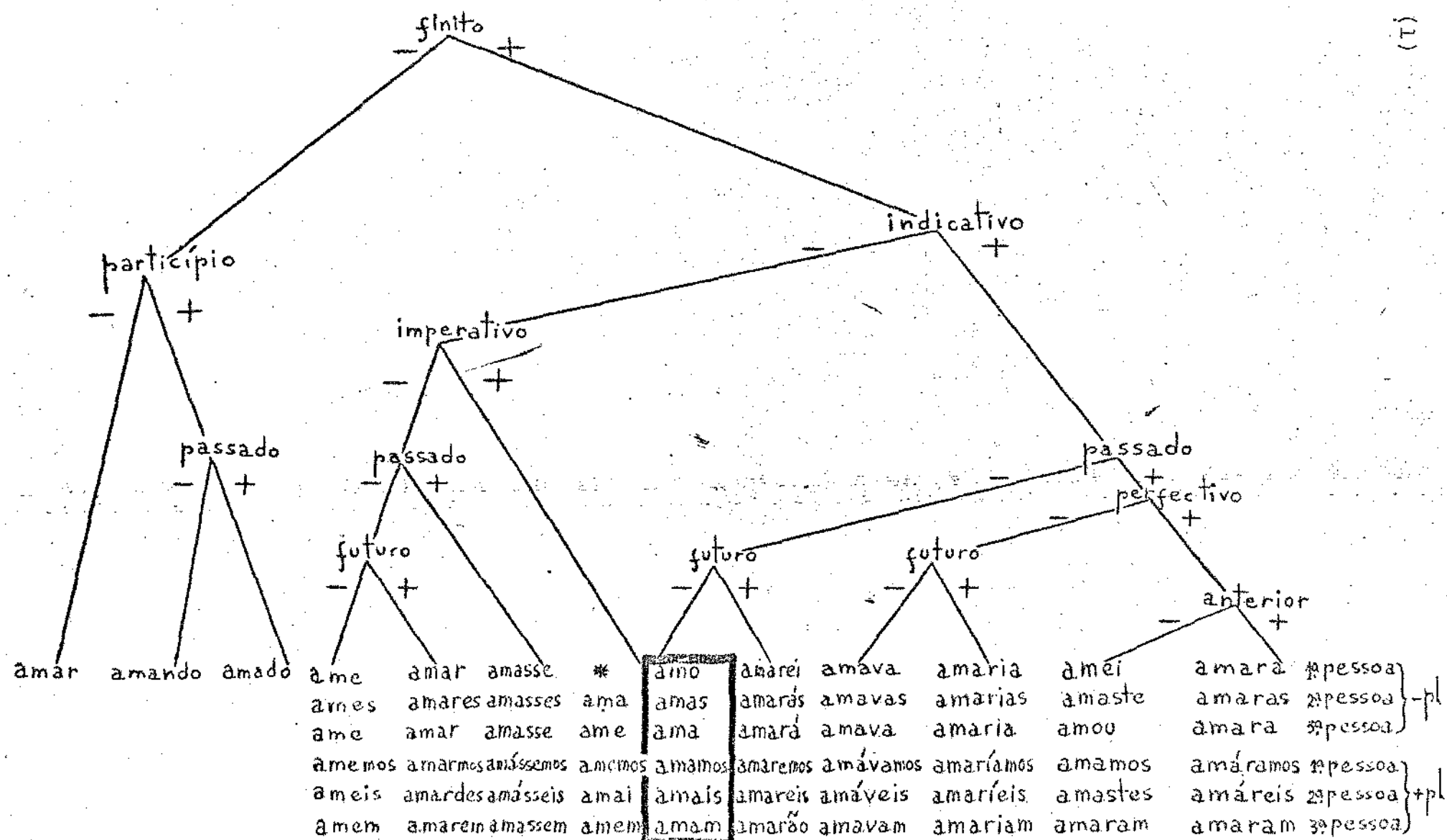
Esta apresentação representa uma simplificação dos fatos mas não nos propomos analisar detalhadamente as várias categorias gramaticais ocorrentes nos verbos. Isto constituiria sem dúvida assunto para um trabalho à parte. Nesta etapa da expo-

sição, o essencial é saber que o grande número de formas verbais da língua resulta das possibilidades de combinação dessas noções gramaticais às bases verbais. As formas podem ser agrupadas em um diagrama como o da pág. 59 (Cf. Harris, 1969 : 63), onde cada uma se encontra perfeitamente distinguida das demais por um complexo de propriedades. Assim, somente a forma amo pode ser caracterizada como:

- | |
|--------------|
| + finito |
| + indicativo |
| - passado |
| - futuro |
| + 1a. pessoa |
| - plural |

Os rótulos empregados na distinção das formas são quase todos parte da terminologia tradicional e referem-se às noções gramaticais de modo, tempo, aspecto, número e pessoa. As formas caracterizadas como [+ finito] são as que admitem flexão de número e pessoa. O fato de que estes rótulos se constituem em um sistema de referências cruzadas, leva a crer que tenham realmente um valor funcional no estudo da sintaxe verbal portuguesa. Com efeito, representam propriedades independentes que podem ser levadas em conta na formulação de regras do componente sintático. Não é nosso propósito fazer um estudo de sintaxe verbal, esse o motivo por que utilizamos a terminologia tradicional. A adequação ou não de tais rótulos só poderá ser testada quando tal estudo tiver sido completado.

Teoricamente, aliás, a análise do componente fonológico de uma língua pressupõe a análise do componente sintático, justa



mente para que se possa trabalhar com propriedades cuja funcionalidade sintática tenha já sido atestada. Na prática, no entanto, não se pode esperar que a análise completa da sintaxe seja feita a fim de que se possa iniciar o estudo da fonologia. Problemas como o que apontamos são comuns, portanto. Em síntese, o que sugerimos aqui é que propriedades independentes podem ser significativas para o estudo da classe verbal. Poderíamos ter utilizado letras ou números, por exemplo, no diagrama, para caracterizar cada uma das formas verbais. ^{contudo} A chamamos ^{que} isto não passaria de um artifício de análise e, feitas as ressalvas acima, não vemos porque não manter a terminologia tradicional em um trabalho como este. O que, em última análise, facilitará a identificação das formas a que se fizer referência no decorrer da análise.¹⁰

3.3 Dados considerados

As formas do presente do indicativo, de que nos ocuparemos principalmente, estão assinaladas em (1). . . . Certas observações se fazem necessárias sobre algumas das formas:

- (a) - procuramos apresentar no diagrama um quadro completo das formas verbais que podem ocorrer na língua. Este o motivo por que não omitimos o chamado pretérito mais-que-perfeito do indicativo. Achamos que o fato de tais formas ocorrerem em determinados dialetos regionais e variantes estilísticas da língua constitui justificativa suficiente para que sejam consideradas em uma análise mais geral;

- (b) - pelo mesmo motivo exposto acima consideramos também a segunda pessoa do plural (P_5 , vós amais). Discordamos nesse ponto de Harris, quando ele apresenta os motivos pelos quais deixou de incluir na análise do espanhol esse complexo número-pessoal:

Second person plural forms (vosotros amáis) occur only marginally in the dialect under investigation. Educated speakers learn them in school, encounter them in literature, and occasionally use them — often incorrectly — in religious and other types of stylized discourse. That is, these forms have in Mexican Spanish roughly the same status as that of "thou" and "ye" forms in standard American English. I therefore see no reason to include them in this study. (1969 : 64)

As observações feitas sobre o uso restrito desta forma em espanhol são sem dúvida aplicáveis ao português. Não obstante, não se deve esquecer que a forma ocorre eventualmente e, por esse motivo, deve ser descrita como as demais. (Além do mais, deixar de lado a segunda pessoa do plural, na descrição do português, seria abandonar um dado importante para a consideração da alternância a/e, na língua - cf. amávamos, amáveis);

- (c) - As formas do chamado infinitivo flexionado, criação do português, não estão incluídas no diagrama. Isso se explica pelo fato de que o próprio termo "infinitivo flexionado" constitui um paradoxo, já que as formas não-finitas (entre as quais se inclui o infinitivo) são, por definição, as que não recebem flexão de número e pessoa. Não faria sentido, portanto, inclu

ir uma ramificação a mais entre as formas marcadas [- finito] para incluir o infinitivo flexionado, que, justamente pelo fato de ser flexionado, deveria ser marcado como [+ finito].

Por constituírem uma idiossincrasia do português, tais formas devem ser tratadas à parte. Não estudaremos aqui as possíveis origens do infinitivo flexionado, nem o seu emprego atual na língua. Este problema, que envolve aspectos extremamente interessantes, deve ser abordado em uma análise do componente sintático;

- (d) - Todas as formas verbais do português são tratadas como cadeias de formativos (formativos lexicais aos quais se acrescentam formativos gramaticais para exprimir as noções de modo, aspecto, tempo, número e pessoas). Tais cadeias são por vezes facilmente segmentáveis. Considere-se por exemplo a forma amávamos, onde se identificam os seguintes formativos:

am + á + va + mos

Esta possibilidade de segmentação do vocábulo verbal em um radical e uma série de sufixos não deixou de ser percebida pelos gramáticos tradicionais. Nos trabalhos estruturalistas dedicados ao assunto, é já apresentada de forma sistemática. Em Camara (1970), por exemplo, a descrição é feita nos seguintes termos:

O verbo é em português o vocábulo flexional por excelência, dadas a complexidade e a multiplicidade das suas flexões. As 2 noções gramaticais de tempo e modo, de um lado, e, de outro lado, de pessoa e número do sujeito, que a forma verbal indica um princípio, correspondem a duas desinências, ou sufixos flexionais, que podemos chamar, respectivamente, sufixo modo-temporal (SMT) e sufixo número-pessoal (SNP). Eles se aglutinam intimamente num global sufixo flexional (SF), que se adjunge ao tema do verbo (T), constituído pelo radical (R) seguido da vogal temática (VT) da conjugação correspondente.

No padrão geral dos verbos portugueses o radical é uma parte invariável. Constituído de um morfema lexical, acrescido, ou não, de um ou mais morfemas derivacionais, ele nos dá a significação lexical, permanente, do verbo. A indicação das noções gramaticais (1 - modo e tempo, 2 - número e pessoa) cabe ao sufixo flexional com seus dois constituintes aglutinados.

Assim, temos uma fórmula geral da estrutura do vocábulo verbal português:

$$T (R + VT) + SF (SMT + SNP)$$

Levando-se em conta a alomorfia de cada um dos sufixos flexionais, e a possibilidade de zero (Ø) para um deles ou ambos, tem-se nesta fórmula a regra geral da constituição morfológica do verbo português. (94)¹¹

Há casos em que a segmentação é evidente. Outros porém apresentam problemas de análise, por não estarem presentes todos os sufixos preditos pela fórmula. Considere-se por exemplo a forma amo (P₁, /presente do indicativo). Por comparação com as demais formas do paradigma conclui-se que am- é o radical. O que dizer, no entanto, a respeito do único sufixo que tem representação fonológica, -o ? Há pelo menos duas possibilidades de análise: (1) considerar -o um alomorfe do sufixo número-pessoal para P₁ - presente do indicativo, e dizer que tanto a vogal temática como o sufixo modo-temporal repre-

sentam-se pelo seu alomorfe zero; (2) considerar que a vogal temática é zero e que o -o carrega cumulativamente as noções de modo-tempo e número-pessoa. A decisão, neste caso, é forçosamente arbitrária, embora análises coerentes possam resultar das duas possibilidades de escolha.

Em uma análise transformacional procura-se a solução que reflita os processos naturais operantes na língua. Assim, em um caso como o de amo, procuraremos mostrar que, a uma forma básica am + a + o (com a vogal temática presente) aplica-se u ma regra fonológica que faz cair a vogal em determinados ambientes. É importante observar que é justamente pelo fato de tal regra ter motivação independente na língua que se pode argumentar a favor do caráter não arbitrário desta análise. Por motivação independente entende-se que a regra que faz cair a vogal não é postulada apenas para dar conta deste caso específico, sendo necessária também para a descrição de outros dados da língua.¹² Será portanto nossa principal preocupação, ao discutir os casos problemáticos do português, a busca de soluções que atendam ao critério de naturalidade de análise.¹³

Passaremos a seguir ao estudo das formas do presente do indicativo no espanhol e no português com vistas à consideração das representações subjacentes e manifestações fonéticas das duas línguas. Lembremos mais uma vez que será nosso objetivo principal identificar sistematicamente os problemas, já que soluções para muitos deles dependem com certeza da reformulação de determinados princípios teóricos.

3.4 Análise dos dados

Ao introduzirmos as formas utilizaremos, para maior facilidade de identificação e referência, a terminologia tradicional. Deve-se ter sempre em mente, no entanto, que cada um dos rótulos mencionados corresponde a um complexo de propriedades sintáticas que distinguem aquele grupo de formas de todos os demais. Assim, quando nos referimos ao presente do indicativo estamos agrupando sob esta denominação todas as formas que podem ser caracterizadas como [+ finito, + indicativo, - passado, - futuro]. É importante que se faça esta observação por que qualquer uma destas propriedades poderá eventualmente ser levada em conta na formulação das regras para os demais complexos temporais.

3.4.1 As formas do presente do indicativo

3.4.1.1 Primeira conjugação

Considerem-se as formas do presente do indicativo da primeira conjugação regular do português, verbo amar:

(2)	P ₁	amo	[ãmu]
	P ₂	amas	[ãmas]
	P ₃	ama	[ãma]
	P ₄	amamos	[ãmãmus]
	P ₅	amais	[ã'mays]
	P ₆	amam	[ãmãw]

Comparem-se estas formas com as formas do espanhol:¹⁴

(3)	P ₁	amo	[amo]
	P ₂	amas	[amas]
	P ₃	ama	[ama]
	P ₄	amamos	[a'mamos]
	P ₆	amam	[aman]

Observe-se que no espanhol as manifestações fonéticas constituem-se em um paradigma mais regular, enquanto que no português há diferenças principalmente com relação ao sistema vocálico. Como teremos oportunidade de comparar, é exatamente no sistema vocálico que reside grande parte das diferenças entre as formas superficiais das duas línguas. As vogais nasais, bem como os ditongos nasais em final de palavra, são um traço característico do português e inexistem no espanhol.¹⁵ Como no português há alternância entre [a] e [ã] nas formas apresentadas em (2), há necessidade de postular uma regra para dar conta da ocorrência das variantes em ambientes específicos. No espanhol não há alternância e portanto tal regra não se justifica. Este é pois um caso em que a diferença entre as duas línguas se explica pelo número de regras aplicadas às formas básicas.

Apresentaremos a seguir a síntese da análise proposta por Harris para as formas em (3) e as regras suplementares necessárias para dar conta do português (2). Em resumo, o que se tem no caso do espanhol é o seguinte:

(a) - radical am;

(b) - vogal temática a que, acrescida ao radical am, constitui o tema da primeira conjugação. A vogal temática não se manifesta foneticamente em P_1 mas a sua presença é reconhecida na representação fonológica básica: [am + a + o], para que se mantenha uma maior regularidade paradigmática. A ausência da VT na representação fonética é explicada pela aplicação

à representação fonológica, de uma regra de truncamen
to:

$$V \longrightarrow \emptyset / + \text{---} + o;$$

- (c) - desinências número-pessoais o, s, Ø, mos, n, acresci-
das ao tema ama. Regras especiais dão conta dos vá-
rios complexos número-pessoais:¹⁶

$$\begin{aligned} \text{a. } [1 \text{ pess}] &\longrightarrow \left\{ \begin{array}{l} o / [- \text{plural}] \\ mos / [+ \text{plural}] \end{array} \right\} \\ \text{b. } [2 \text{ pess}] &\longrightarrow s \\ \text{c. } [3 \text{ pess}] &\longrightarrow n / [+ \text{plural}] ; \end{aligned}$$

- (d) - acento na penúltima vogal das manifestações fonéticas.

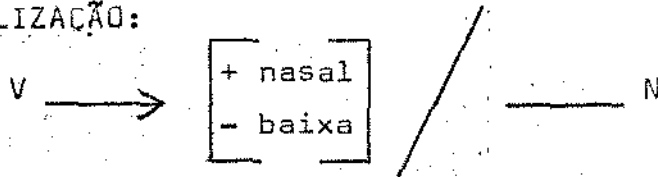
Uma primeira versão da regra de acento para as formas
verbaís, ordenada ^{após} a regra de truncamento, dá conta do
padrão acentual nas formas do presente do indicativo:

$$V \longrightarrow [1 \text{ acento}] / \text{---} C. V C. \#]_{\checkmark}$$

Passemos agora à consideração das formas do português. Ob-
serve-se que a vogal do radical é [ã]. Considerando-se ape-
nas as diversas ocorrências do radical [ãm-], poder-se-ia pen-
sar que a vogal subjacente é [ã]. No entanto, um breve exame
de maior número de dados da língua é suficiente para mostrar
que [ã] e [a] estão sempre em alternância, ocorrendo [ã] dian-
te de qualquer segmento nasal. Esta alternância, com efeito,
é mais geral no português, atingindo todo o sistema vocálico.
Pode-se dizer que há na língua um processo geral de nasaliza-
ção de vogais diante de um segmento nasal e que as vogais na-
salizadas são sempre [- baixas]. Isto significa que as vo-
gais subjacentes [i], [u], [o] e [e] realizam-se como [ĩ] ,

[ũ], [õ] e [ẽ] diante de segmento nasal, e [a], [ɛ] e [o] realizam-se como [ã], [ẽ] e [õ], respectivamente, neste mesmo ambiente. Esta é a origem das chamadas vogais nasalizadas do português.¹⁷ No caso amar parte-se da forma subjacente [am-] para o radical e explica-se a ocorrência de [õ] como resultado da aplicação, a essa forma, da regra seguinte:

(4) NASALIZAÇÃO:



(onde V representa qualquer vogal e N qualquer segmento nasal).

Justifica-se internamente, portanto, a escolha de [am-] como representação fonológica do radical de, amar em português, o que corresponde, em última análise, ao radical identificado no caso do espanhol.

Ao radical am- segue-se a vogal temática a. Também no português esta vogal não se manifesta em P_1 , mas reconheça-se sua presença na representação fonológica para manutenção da regularidade no plano paradigmático. A regra de truncamento postulada para o espanhol, que faz cair a vogal em determinados ambientes, aplica-se também ao português. Exemplos da aplicação dessa mesma regra fora da classe verbal fornecem motivação independente para sua existência na língua. Assim, quando a base nominal como [menino] é acrescentado o sufixo flexional de gênero feminino [-a], -a última vogal da base é truncada, por efeito dessa mesma regra.¹⁸ Apresentamos aqui a regra de truncamento em uma forma mais geral, para dar conta da queda de vogais antes de outras vogais que não a:

(5) TRUNCAMENTO



(A generalização do ambiente para.../ + + V se evidencia quando da consideração das formas do presente do subjuntivo, mas a limitação de tempo não nos permite considerar o assunto neste trabalho).

Uma vez justificada a presença da vogal temática na representação fonológica de todas as formas, resta explicar a alternância [a] ~ [ã]. A alternante que ocorre em P₄ é resultado da aplicação da regra de nasalização.⁴

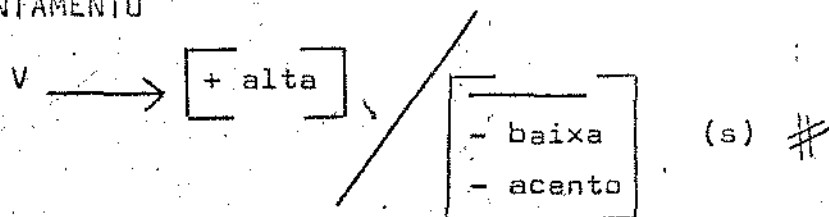
Temos, portanto, como representação fonológica do tema de amar, o radical | am - | e a vogal temática | -a- |: | am + a|. Este tema, como veremos, está presente em todas as formas que constituem o paradigma deste verbo.

Consideremos a seguir as desinências número-pessoais. Em P₁ identifica-se [u] como marca de número-pessoa, enquanto que no espanhol tem-se [o]. Evidências internas levam-nos a postular a vogal subjacente |o| também no caso do português. Sabe-se que o quadro de vogais que ocorrem em posição átona final reduz-se a três: [u], [i] e [a]. (Cf. Camara, 1970 : 34). Pode-se, portanto, dizer que uma regra de levantamento das vogais [o] e [e] aplica-se neste ambiente. Observe-se que o levantamento ocorre também diante de [s], segundo se depreende dos exemplos seguintes que representam nomes em sua forma singular e plural (estes acrescidos do sufixo flexional de número, -s):

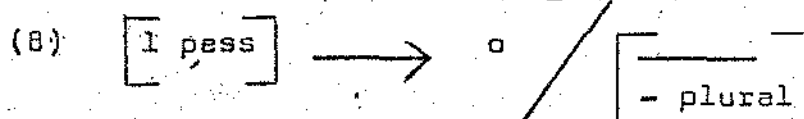
(6)	(sg)		(pl)		
	leite	[<u>'leyt'i</u>]	-	leites	[<u>'leyt'is</u>]
	dado	[<u>'dadu</u>]	-	dados	[<u>'dadus</u>]

A regra de levantamento é a seguinte:

(7) LEVANTAMENTO



Tem-se, portanto, para P_1 , como em espanhol:



Note-se que o português difere do espanhol pela aplicação à vogal átona final da regra de levantamento (7). Observe-se também que esta é uma das regras mais tardias, havendo mesmo um estágio na derivação das regras, antes da aplicação de (6), em que não ocorrem na língua vogais altas em posição final não-acentuada. Trataremos deste problema ao estudar as formas da terceira conjugação a seguir. Por ora basta dizer que esta pode ser considerada uma regra de detalhe fonético, na língua, havendo dialetos aos quais ela ^{não} se aplica. Em outros dialetos a regra é aplicada em uma forma menos geral, isto é, ocorre o levantamento da vogal $\left[\begin{array}{l} - \text{ baixa} \\ + \text{ posterior} \end{array} \right]$ (i.e., o) apenas, neste ambiente (Cf. Naro, 1971). No dialeto em que nos baseamos a regra se aplica em sua forma mais geral.

Para P_2 e P_3 tem-se no português $[s]$ e $\emptyset$, respectivamente, também como no espanhol. Já que a marca número-pessoal é fonologicamente nula em P_3 , necessita-se apenas da regra para P_2 :

$$(9) \quad [2 \text{ pess}] \longrightarrow s \quad / \quad [- \text{plural}]$$

(No caso do português faz-se necessário especificar o ambiente para esta regra, já que estaremos também considerando P_5 , ou seja, a segunda pessoa do plural).

Em P_4 , tem-se [mus] como manifestação da marca de número-pessoa. Pelos mesmos motivos apresentados quando da consideração de P_1 , acreditamos que a representação fonológica seja | -mos |. A ela se aplica a regra (6), de levantamento. Mantém-se pois, até o momento, o paralelismo entre o português e o espanhol, com respeito às formas subjacentes para as marcas número-pessoais. Para P_4 tem-se, também no português:

$$(10) \quad [1 \text{ pess}] \longrightarrow mos \quad / \quad [+ \text{plural}]$$

Ao estudarmos P_5 , a seguir, não faremos referência à análise de Harris, já que esta forma não é por ele considerada. Achamos no entanto que as observações que fazemos com relação ao português são também aplicáveis ao espanhol, uma vez que, segundo o próprio Harris, a forma ocorre "marginalmente" no dialeto por ele analisado, como vosotros amáis. Tem-se assim uma manifestação fonética idêntica a P_5 em português.

Iniciaremos a nossa consideração de P_5 pelo exame de algumas formas verbais da língua que manifestam as noções 2a. pessoa e plural. Apresentaremos aqui formas das três conjugações, já que o que nos interessa no momento é a identificação do sufixo flexional número-pessoal:

(11) I

amais [ã'mays]
 comeis [kõ 'meys]
 unis [ũ'niys]

II

credes [kred'is]
 ledes [led'is]
 vedes [ved'is]
 ides [id'is]
 rides [hid'is]
 vindes [vid'is]
 podes [põd'is]

(11) I constitui uma lista aberta, dela fazendo parte a 2a. pessoa do plural da maioria absoluta dos verbos da língua (inclusive dos que constantemente se formam na 1a. conjugação).

Identifica-se a marca número-pessoal [ys] para estas formas.

(11) II é uma lista fechada. Dela fazem parte as exceções de I, um pequeno grupo de verbos de infinitivo foneticamente monossilábico¹⁹ (crer, ler, ver, ir, rir, vir, pôr) que apresentam [d'is] como marca número-pessoal. Estamos, portanto, diante de um exemplo de alternância ou de supleção. A tarefa do analista, em um caso como este, resume-se no seguinte: procurar evidências na própria língua que permitam postular uma representação subjacente à qual se aplicam regras para dar conta da alternância. Na falta de evidência é forçoso reconhecer que se trata de um caso de supleção.

Voltemos aos exemplos do português. Para melhor compreensão do problema é importante acrescentar aqui mais um dado significativo: P₅, nos verbos que fazem parte de I, manifesta-se como [ys] em todos os tempos, à exceção do futuro do subjuntivo e do chamado infinitivo flexionado (onde ocorre [d'is] (a mardes)), e do pretérito perfeito do indicativo (onde ocorre

[st'is] (amastes)). Há argumentos que mostram que as marcas número-pessoais do pretérito perfeito do indicativo devem ser consideradas isoladamente, tratando-se, portanto, de um caso de supleção. Não há, no entanto, argumentos convincentes no caso de P_5 no futuro do subjuntivo e infinitivo flexionado. Esses exemplos, pelo contrário, parecem reforçar a hipótese de que em P_5 há alternância entre [ys] e [d'is], devendo ser postulada uma forma básica e regras que dêem conta das duas manifestações em ambientes específicos. Uma vez reconhecida a alternância, o problema é a escolha da forma básica que atenda ao critério de naturalidade na análise da língua. Examinaremos duas hipóteses:

- (a) - considerar [-is] como forma básica.²⁰ Nesse caso seria necessária uma regra para inserção de d nos casos em que ocorre [d'is] (em amardes e em P_5 de todos os verbos monossilábicos da lista (11) II). Nos casos em que ocorre [ys] seria necessária uma regra para inverter a vocalicidade do primeiro segmento, e que é responsável pela ocorrência de semivogais em determinados ambientes;
- (b) - considerar [-des] como forma básica e postular inicialmente uma regra que faz cair o d em posição intervocálica. Após a queda do d aplicar-se-ia à vogal final não acentuada precedida de s a regra de levantamento já apresentada anteriormente (7)²¹. Obter-se-ia então [is], forma à que se aplicaria por último a regra que converte vogais em semivogais. Nes

se caso as formas de (11) II teriam que vir marcadas no léxico como não sujeitas à aplicação da regra de queda do d.

Examinemos, a seguir, as duas hipóteses. Inicialmente mencionamos a possibilidade de partir da forma básica | -is | e postular uma regra de inserção do d em amardes e em P_5 nos verbos em II. Tal regra nos parece totalmente ad hoc, já que aplicável apenas a esses casos e, portanto, sem motivação independente na língua.

Passemos agora à consideração da segunda hipótese, segundo a qual deve-se partir da forma básica | -des | e postular uma regra de queda do d. Ao contrário da regra de inserção de d, a regra de queda de d é necessária para uma descrição sincrônica da língua, conforme se deduz do exemplo seguinte (onde há alternância entre presença e ausência deste segmento em formas relacionadas):

- | | | |
|------|----|----------|
| (12) | I | II |
| | pé | pedestre |
| | | pedestal |

Observe-se que há um denominador comum semântico entre as formas em (12) II e a forma em (12) I, apesar da diferença na configuração fonológica do formativo lexical em questão. Em todos os casos o que se tem é a mesma forma básica |pedest|, sendo que em (12) I a forma é marcada como [+E] e a ela se aplica a regra de queda de d intervocálico, formulada a seguir:

- (13) QUEDA DE d.²² (aplicável a formativos [+E])

$d \longrightarrow \emptyset \quad / \quad V \text{ ______ } V$

Tem-se portanto:

	p&de #
1. acento	pêde #
2. queda de <u>d</u>	pê e #
3. assimilação	pê & #
4. crase	pê #
5. redução	pê #

(as regras de acento, assimilação, crase e redução não se rão discutidas em detalhe neste trabalho, uma vez que não são relevantes para o assunto de que aqui nos ocupamos).

Uma vez comprovado o fato de que a regra de queda de d in tervocálico é necessária na língua para explicar dados fora da classe verbal, é natural que ela se aplique também às formas verbais. Parece-nos, portanto, mais natural, a segunda hipóte se por nós considerada com relação à caracterização da marca número-pessoal em P_5 . Partindo-se então desta hipótese, o que se tem é o seguinte:

(a) - formativo flexional número-pessoal para P_5 | -des |.

Como os demais formativos flexionais, | -des | é con siderado [+E], a ele se aplicando todas as regras que atingem os formativos constituintes do estreto [+E] da língua (Cf. capítulo 2). Tem-se, portanto , para P_5 do presente do indicativo de amar, por exem plo, | am + a + des# |;

(b) - aplicação da regra (13), queda de d:

	am+a+des #
1. acento	am+á+des #
2. queda de <u>d</u>	am+á+ es #

Observe-se que a regra se aplica ao formativo flexional, que é sempre [+E], mesmo que o formativo lexical seja [-E] (Cf. 2.3). Assim, no caso de ceder, onde o formativo lexical ced- é [-E] e portanto o d é mantido em todas as formas, tem-se, para P_5 , presente do indicativo:

| sed] [-E] +e + des] [+E] †

1. acento sed + é + des †

2. queda de d sed + é + es †

(c) - aplicação da regra (7), levantamento:

| am+a+des †

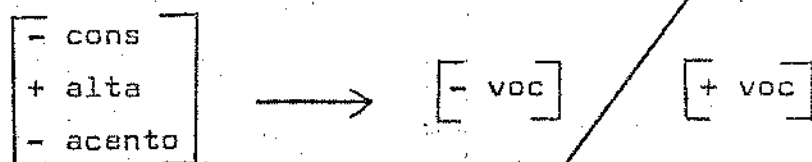
1. acento am+á+des †

2. queda de d am+á+ es †

3. levantamento am+á+ is †;

(d) - aplicação de uma regra de formação de semivogais. Tal regra, prevista nas duas hipóteses, é uma regra geral do português. Sempre que uma vogal alta não acentuada precede ou segue outra vogal, ela adquire o traço não vocálico, isto é, transforma-se em semivogal. Esta é a origem dos vários ditongos crescentes e decrecentes existentes na língua. A regra em questão é a seguinte:

(14) FORMAÇÃO DE SEMIVOGAL ²³



Explica-se, então, a forma [õ'mays] :

	am+á+des ‡
1. acento	am+á+des #
2. queda de <u>d</u>	am+á+ es ‡
3. levantamento	am+á+ is ‡
4. formação de semivogal	am+á+ ys ‡
5. nasalização	ãm+á+ ys ‡

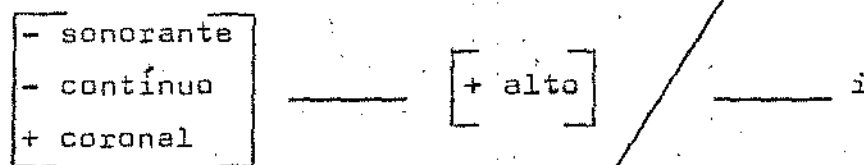
Admitindo-se que a segunda hipótese é a mais natural, resta explicar os casos em que ocorre [d'is] e não [ys] como marca de P₅. Estes casos são os seguintes:

(a). - P₅ no futuro do subjuntivo e infinitivo flexionado.

Em ambos os casos explica-se automaticamente a forma amardes a partir de | am+á+re+des # |, uma vez que a regra de queda de d (13) aplica-se apenas entre vogais, e no momento de sua aplicação a esta forma a vogal que precede o d já caiu por efeito de uma regra de apócope (a ser vista posteriormente. Cf. (29)).²⁴

Não havendo condições para a aplicação da regra neste caso, mantém-se o d que posteriormente se palataliza diante de i, por uma regra geral deste dialeto do português. A regra de palatalização é a seguinte:

(15) PALATALIZAÇÃO



(isto é, t e d palatalizam-se diante de i).²⁵

A motivação para a regra (15) é fornecida por exemplos como os seguintes, onde se verifica que há

alternância entre [t, d] e [t', d'] :

(16)	I	II
mata	[ˈmata]	
matei	[maˈtey]	mate [ˈmat'i]
matou	[maˈtow]	
roda	[ˈhoda]	
rodei	[hoˈdey]	rode [ˈhod'i]
rodou	[hoˈdow]	

Um breve exame fora da classe verbal mostra que a regra de palatalização aplica-se a todos os itens da língua, [+ē] ou [-ē], sempre que há ambiente, isto é, sempre que t ou d ocorrem diante de i. Alguns exemplos são:

tia	[t'iya]	
dia	[d'iya]	
diabo	[d'i'abu]	
feitizo	[fey't'isu]	, etc.

Observando-se a forma amardes [ã'mard'is] conclui-se que, sendo [-des] a representação fonológica para número-pessoa em P₅, a regra de palatalização (15) deve ser ordenada após a regra de levantamento (7), já que a última fornece ambiente ([i]) para aplicação da primeira.

- (b)- P₅ nos verbos apresentados em (11) II (credes, ledes, etc.). Em todas as formas de (11) II o d encontra-se entre vogais e seria de se esperar que a regra (13), queda de d, se aplicasse. Não são gramaticais, no entanto, formas como * [leys], * [kreys], etc., para P₅. Como (11) II constitui uma lista fechada da qual faz parte um pequeno número de verbos, pode-se pensar em marcar es-

tes radicais no léxico como condicionadores da ocorrência de [d'is] para marca de número-pessoa em P_5 . Esses casos seriam tratados como exceções, portanto.

Concluimos aqui as considerações sobre a marca de número-pessoa em P_5 , no português, estabelecendo que:

$$(17) \quad [2 \text{ pess}] \longrightarrow \text{des} \quad / \quad \boxed{+ \text{ plural}}$$

Passemos agora ao estudo de P_6 . Nesta forma o que se tem foneticamente, no caso de amar, é [w̃] como marca número-pessoal. No espanhol, como já vimos, esta marca se manifesta foneticamente como [n], identificando-se sem problemas [-n] como representação fonológica. No português a identificação não é tão simples. Em primeiro lugar, observa-se que o que se tem na superfície é uma semivogal nasalizada. Estabelecer tal segmento como subjacente nos parece fora de cogitação, porque tanto as chamadas semivogais quanto os segmentos nasalizados que ocorrem na língua parecem ser produto da aplicação de regras fonológicas a certos segmentos subjacentes em determinados ambientes. Além disso, a semivogal nasalizada [w̃] constitui, juntamente com a vogal temática (que se manifesta em P_6 - 1ª. conjugação - como [ã]), o ditongo decrescente [ãw̃], comum no português em final de palavra, e que vem a ser mesmo uma das características marcantes desta língua em confronto com as demais línguas românicas. Procuraremos demonstrar que no português, como no espanhol, o que se tem, como representação fonológica, é |am+a+n|.

Considerem-se inicialmente as seguintes séries de exemplos:

(18)	I		II
	pão	[...ãw]	panificação
	mão		manipular
	irmão		irmanar
	leão		leonino
	capitão		capitanear

Não há dúvida de que existe um denominador comum semântico, entre as formas de (18) I e as correspondentes de (18) II. As formas em (18) I são [+E], enquanto que as de (18) II são [-E]. Note-se a alternância entre o ditongo nasal em (18) I, e o segmento n em (18) II. A existência da alternância permite que sejam postuladas para as formas em (18) I as seguintes representações subjacentes:

(19)	pane #
	mano #
	irmano #
	leone #
	kapitane #

Nas representações sistemáticas a escolha da vogal final é feita com base nas formas do plural, onde se tem, respectivamente: pães, mãos, irmãos, leões, capitães.²⁶ Ou seja:

(20)	pane + s #
	mano + s #
	irmano + s #
	leone + s #
	kapitane + s #

Reconhecer que certos formativos lexicais apresentam a vogal final e nas representações subjacentes para o singular implica em uma simplificação de pelo menos duas regras da língua: formação de plural (que passa a ser representado simples

mente pelo formativo gramatical $[-s]$ acrescido às bases nominais terminadas em $-a$, $-e$ ou $-o$) e acentuação (explicando-se a ocorrência superficial de um número considerável de oxítonos que parecem contrariar a regra geral de acentuação da língua, segundo a qual o acento recai na maioria dos casos na penúltima sílaba.²⁷ Tais oxítonos^{são} na verdade paroxítonos por ocasião da regra de acento, aplicando-se depois uma regra de epócope que faz cair a vogal e em determinados ambientes. Voltaremos a esta regra oportunamente). No momento o que é importante destacar é a presença do segmento n nas representações subjacentes deste grupo de nomes que apresentam foneticamente ditongos nasais no singular e no plural. Isso nos permite dizer que em P_6 está também presente este segmento nas representações fonológicas, sendo ele o responsável pelos ditongos nasais ocorrentes na superfície. Admitida tal hipótese, evidencia-se mais uma vez a semelhança de formas básicas para o espanhol e o português. Nas duas línguas, a representação subjacente seria, no caso, $|am+a+n\#|$, aplicando-se no caso do português um maior número de regras que dariam conta finalmente de $[\tilde{a}m\tilde{a}w]$. Vejamos quais seriam estas regras. Inicialmente, a uma representação como $|am+a+m\#|$, aplica-se a regra (4) de nasalização de vogal que precede um segmento nasal. O resultado seria, no caso, $\tilde{a}m+\tilde{a}+n$ (já que as vogais nasalizadas são, como vimos, $[-baixas]$). O problema que imediatamente se coloca é a origem do ditongo nasal ocorrente na superfície. Apresentaremos a seguir uma possível explicação para esses ditongos.

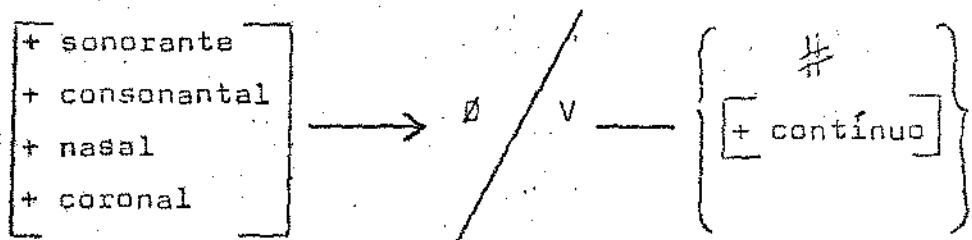
Examinemos inicialmente o comportamento do segmento n nos grupos de exemplos que se seguem:

(21)	I	II
a. (18)	pão	panificação
	mão	manipular
	irmão	irmanar
	leão	leonino
	capitão	capitanear
b.	lua	lunar
	areia	arenoso
	moeda	monetário
	veia	venoso
	doação	donativo

Como vimos antes, as formas em (21) II, existentes no português atual, levam-nos a postular para as formas em (21) I representações subjacentes onde está presente o segmento n. Exemplos seriam: (21) a. |pané#|, |mano#|; (21) b. |luna#|, |arena#|. ²⁸ Para explicar as formas de (21) I é necessário então postular uma regra que faça cair o segmento nasal (após ter causado nasalização da vogal precedente por efeito de (4)). ²⁹ Esta regra faz parte do conjunto de regras que se aplicam apenas aos formativos da língua especificados como [+E] (como todos os formativos em (21) I). Por outro lado, a regra de nasalização (4), que deve ser ordenada antes da regra que faz cair o segmento nasal em certos ambientes, é mais geral, aplicando-se sempre que há ambiente, aos formativos [+E] e [-E].

Passemos então à consideração da regra de queda de n, que formularemos como:

(22) QUEDA DE n (aplicável a formativos [+E]).



(i.e., n cai em final de palavra, ou depois de vogal e qualquer segmento especificado como [+ contínuo]).³⁰

Vejamos agora porque é necessário dizer que o segmento na sal cai em final de palavra. Examinemos mais uma vez os dados com os quais vimos trabalhando, mas fazendo desta vez uma nova divisão de acordo com a vogal final das bases nominais:

(23)

I

mão - |mano#| (cf. pl. maos-
|mano+s#|)

irmão - |irmano#| (cf. pl. irmãos-
|irmano+s#|)

II

pão - |pane#| (cf. pl. pães-
|pane+s#|) ³¹

leão - |leone#| (cf. pl. leões-
|leone+s#|)

capitão - |kapitane#| (cf. pl. capitães-
|kapitane+s#|).

Se dissermos que a regra de queda de n entre vogais se aplica tanto às formas de I como às de II, obteremos, para o singular, após a aplicação de (4) -nasalização e (22) -queda de n:

(24)

I

mão

irmão

II

põe

leõe

kapitõe

A essas formas seriam então aplicadas as regras (7) -levantamento, (14) -formação de semivogais e ainda uma regra de nasalização secundária, segundo a qual nasaliza-se qualquer segmento [-consonantal] (i.e., vogais e semivogais) que precede ou se

que uma vogal nasalizada.³² Esta regra é a seguinte:

(25) NASALIZAÇÃO SECUNDÁRIA:

$$\left[\begin{array}{c} - \text{consonantal} \\ \left\langle \begin{array}{c} \text{sonorante} \end{array} \right\rangle \end{array} \right] \longrightarrow \left[\begin{array}{c} + \text{nasal} \\ \left\langle \begin{array}{c} \text{baixa} \end{array} \right\rangle \end{array} \right] / \left[\begin{array}{c} \text{V} \\ + \text{nasal} \end{array} \right]$$

Após a aplicação destas regras obtém-se, então:

(26)

I
 $\left[\begin{array}{c} \text{m} \tilde{\text{o}} \tilde{\text{w}} \\ \text{ir}' \text{m} \tilde{\text{e}} \tilde{\text{w}} \end{array} \right]$

II
 * $\left[\begin{array}{c} \text{p} \tilde{\text{õ}} \tilde{\text{y}} \\ \text{li} \tilde{\text{o}} \tilde{\text{y}} \\ \text{kapi}' \text{t} \tilde{\text{õ}} \tilde{\text{y}} \end{array} \right]$

As formas obtidas em (25) II não ocorrem, no entanto, na língua. Não se pode, pois, aplicar a regra (22) às representações sistemáticas em (23) I e II sem que antes se aplique u ma outra regra às formas de (23) II, que faz cair o e final e impede assim o levantamento desta vogal e posterior transformação em semivogal, após a queda do n. A regra em questão é uma regra de apócope, para a qual há bastante motivação sincrônica. Senão, vejamos:

— reconhecer um e final subjacente em quase todos os nomes que terminam foneticamente em consoante (e. g. flor, mar, nariz, mal, etc.), significa simplificar a regra de formação do plural no português, uma vez que a marca de número plural passa a ser simplesmente $[-s]$. As representações fonológicas seriam, para o singular, respectivamente: $[\text{flore}\#]$, $[\text{mare}\#]$, $[\text{narize}\#]$, $[\text{male}\#]$;

— nos infinitivos, foneticamente oxítonos (e. g. amar^(r), comer^(r), unir^(r)), o reconhecimento do e final nas formas subjacentes implica em uma maior simplificação no tratamento da acentuação, reforçando-se a hipótese de que é paroxítona a grande

maioria dos formativos não monossilábicos da língua.³³ As representações subjacentes para estes infinitivos seriam, respectivamente: |am+a+re#|, |kom+e+re#|, |un+i+re#|;

— Itens estrangeiros que se incorporam ao léxico como empréstimos recebem sempre a vogal final e que é, inclusive, representada ortograficamente (e. g. ingl. 'club' - port. clube; ingl. 'team' - port. time, etc.).³⁴ Isto indica que o falante nativo de certa forma rejeita como estrangeiras as palavras terminadas por consoantes, adaptando-as ao sistema fonológico do português (no caso em questão, introduzindo a vogal e).

Resta identificar então os ambientes em que ocorre a regra de apócope. Examinemos os exemplos seguintes:

(27)	flor	flore#	flores	
	amor	amôre#	amores	
	mal	male#	males	} 35
	sol	sole#	sóis	
	nariz	narize#	narizes	
	mês	meze#	meses	
	pão	pane#	pães	
	cão	kane#	cães	

De acordo com estes dados, verifica-se a queda da vogal após os segmentos r, l, z e n (cf. representações fonológicas).

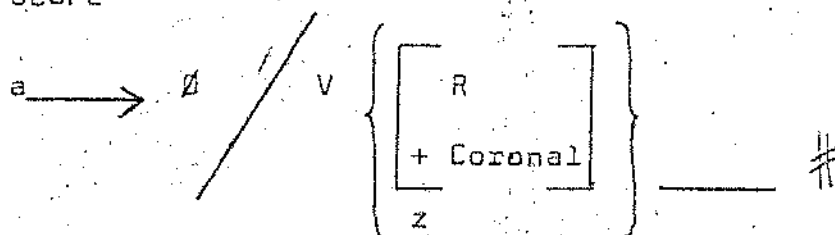
Há casos, no entanto, em que a vogal se mantém, mesmo vindo depois de alguns destes segmentos:

(28)	catre	katre#	catres
	abutre	abutre#	abutres
	rifle	rifle#	rifles

Com base nestes dados, podemos formular como primeira a-

proximação da regra de apócope:

(29) APÓCOPE ³⁶



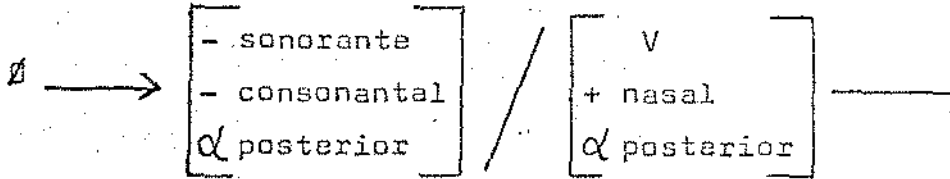
(Onde R abrevia "ressonantes". A regra diz que cai o e final quando é precedido por V (ogal) seguida de r, l, n ou z).

Uma vez formulada a regra, vejamos de que modo a sua aplicação às formas consideradas em (23) II pode levar a resultados diversos dos obtidos em (26) II. Inicialmente, a regra de apócope teria que ser ordenada antes de (22) -queda de n, já que o n pode fornecer ambientes para a sua aplicação. (29) aplica-se a todas as formas em (23) II, podendo ser ordenada antes de (4) -nasalização e (22). -queda de n. O resultado da aplicação desta regra a (23) I e II seria, então, para mão e pão:

(30)	I	II
	mano#	pane#
(29).	—	pan
(4).	mão	pãn
(22).	mão	pã

Atingido este estágio da derivação nota-se que são já conhecidas as demais regras a serem aplicadas a (27) I para que se obtenha o produto fonético final ['mõw]. Estas regras são: (7) -levantamento, (14) -formação de semivogal e (25) -nasalização secundária. No caso de (27) II, entretanto, é necessário postular uma nova regra segundo a qual uma semivogal homorgânica é inserida após uma vogal nasalizada em posição final. ³⁷ Esta regra é a seguinte:

(31) INSERÇÃO DE SEMIVOGAL



A razão pela qual indicamos na regra que a semivogal inserida concorda em posterioridade com a vogal nasalizada é a existência de exemplos como:

(32)

- I. bem |bene#| (cf. benévolência) - ['bẽy]
comem |kom+e+n#| - ['kõmẽy]
sentem |sent+e+n#| - ['sẽntẽy]

e, possivelmente, pelo menos na variante estudada:³⁸

- II. fim |fin#| - ['fĩy]
atun |atun#| - [a'tũw]

(cf. pl. fins e atuns - não há evidência para e nas representações subjacentes.

Os chamados ditongos nasais que caracterizam o português podem então ser explicados como sendo o resultado da aplicação, às formas subjacentes, das regras que se seguem:

- (33) I (representações subjacentes com e final) II (representações subjacentes com o final)

(26). apócope(4). nasalização(4). nasalização(22). queda de n(22). queda de n(7). levantamento(28). inserção de semivogal(14). formação de semivogal(25). nasalização secundária(25). nasalização secundária

No caso das formas de P_6 , como no espanhol, o que se tem

como marca número-pessoal é [-n]. Pode-se então dizer que:

- (34) $\left[\begin{array}{l} 3 \text{ pess} \end{array} \right] \longrightarrow n / \left[\begin{array}{l} + \text{plural} \end{array} \right]$

Observe-se que no caso das formas verbais as únicas regras que se aplicam e que dão conta dos ditongos finais são :

(4) -nasalização, (22) -queda de n, (28) -inserção de semi-
e (25) -nasalização secundária
vogal ✓ja que não há vogal final nas formas subjacentes.³⁹

Têm-se então, a partir da consideração das formas do presente do indicativo dos verbos da Primeira Conjugação, os seguintes formativos gramaticais como marcas de número-pessoa :

(35) FORMATIVOS GRAMATICAIS, NÚMERO-PESSOAIS

a.	[1 pess]	→	$\left\{ \begin{array}{l} o \\ mos \end{array} \right. \begin{array}{l} / \\ / \end{array} \left\{ \begin{array}{l} \boxed{- \text{ plural}} \\ \boxed{+ \text{ plural}} \end{array} \right.$
b.	[2 pess]	→	$\left\{ \begin{array}{l} s \\ des \end{array} \right. \begin{array}{l} / \\ / \end{array} \left\{ \begin{array}{l} \boxed{- \text{ plural}} \\ \boxed{+ \text{ plural}} \end{array} \right.$
c.	[3 pess]	→	$n \quad / \quad \boxed{+ \text{ plural}}$

Examinemos a seguir o padrão acentual nas formas que vimos considerando. As formas de (2), como se observa, recebem acento na penúltima sílaba, nas manifestações fonéticas, à exceção de P_5 que apresenta o acento na última sílaba. No entanto, se postularmos uma regra de acento que faz com que ele incida na penúltima vogal das representações subjacentes, veremos que o padrão regular pode ser mantido, uma vez que, como vimos ao estudar a marca número-pessoal de P_5 , a representação subjacente para esta forma é $|am+des\#|$. Partiremos então do princípio de que a regra de atribuição do acento deve ser ordenada antes da regra (13) -queda de d (e, consequentemente antes das regras (7) -levantamento e (14) -formação de semivogal, que implicam em mudança de estrutura silábica nas formas superficiais).

Pode-se então formular a regra de acento para estas for-

mas como:

(36) ACENTO (Formas verbais)⁴⁰

$V \longrightarrow \left[\text{acento } 1 \right] / \text{---} \text{C.VC}^1 \# \right]_{\checkmark}$

(isto é, recebe acento primário, nas formas verbais, a penúltima vogal das representações fonológicas).

É necessário observar ainda que (33) é ordenada depois da regra (5) -truncamento. Considere-se por exemplo P_1 , cuja representação subjacente é $|am+a+o\#|$. Se a regra de acento fosse ordenada antes da regra de truncamento, o acento recairia exatamente na vogal que deve ser truncada, ou seja: $|am+\hat{a}+o\#|$.

A ordenação a ser mantida deve ser, portanto:

(5). -truncamento

(33). -acento

As representações fonológicas coincidem, deste modo, no espanhol e no português, para as formas do presente do indicativo da 1a. conjugação regular. As diferenças superficiais explicam-se, como vimos, por regras posteriores aplicadas às formas do português.

Passemos agora à consideração do presente do indicativo na 2a. e 3a. conjugações.

3.4.1.2 Segunda conjugação

Formas do presente do indicativo da segunda conjugação regular do português, verbo comer:

(37)	P_1	como	$[k\ddot{o}mu]$
	P_2	comes	$[k\ddot{o}mis]$
	P_3	come	$[k\ddot{o}mi]$
	P_4	comemos	$[k\ddot{o}'m\ddot{e}mus]$

P ₅	comeis	[kõ'meyis]
P ₆	comem	[kõ'mẽy]

No espanhol as formas são:

(38) P ₁	como	[komo]
P ₂	comes	[komes]
P ₃	come	[kome]
P ₄	comemos	[ko'memos]
P ₆	comem	[komen]

Note-se mais uma vez que aparentemente apresentam maior regularidade as formas do espanhol. A identificação dos formativos não apresenta problemas:

The root in com; the theme vowel of the second conjugation is e (by definition). Otherwise the present tense of the second conjugation is just like that of the first conjugation (Harris, 1969: 68).

Verifica-se, com efeito, que as marcas número-pessoais que se manifestam são as mesmas já identificadas para as formas da primeira conjugação.

No português a identificação dos formativos parece, à primeira vista, mais problemática. O radical apresenta uma vogal nasalizada ([kõm-]), mas esta nasalização é resultado, como já vimos, da regra (4). A representação fonológica básica do radical pode ser então dada como |kom-|.

Como vogal temática característica da segunda conjugação estabelece-se |-e-| com base na vogal que se manifesta nas formas onde sobre ela recai o acento (P₄ - [kõ'mẽmus] e P₅ - [kõ'meyis]). A nasalização da VT, em P₄, explica-se pela regra (4). Quanto à vogal que ocorre nas demais formas ([i]),

em P_2 e P_3), é o resultado da regra (7) -levantamento. Note-se que não se dá o levantamento da vogal átona em P_6 $[kóme+n]$ porque vem seguida de um segmento outro que não $[-s]$.

Temos, portanto, como representação subjacente do tema de comer, $[kom+e\#]$. Este tema ocorre em todas as formas que constituem o paradigma deste verbo.

As marcas número-pessoais não apresentam nenhum problema especial na 2ª. conjugação, sendo aplicáveis, no caso, as mesmas regras de inserção propostas em (35). Mantém-se também o mesmo padrão acentual previsto em (36) para as formas da 1ª. conjugação.

Verifica-se pois que, também no caso do presente do indicativo da 2ª. conjugação, pode-se partir das representações básicas idênticas para o português e o espanhol, sendo as diferenças que se manifestam nas representações fonéticas atribuídas mais uma vez às regras aplicáveis às formas do português, apenas.

3.4.1.3 Terceira conjugação

Formas do presente do indicativo da 3ª. conjugação regular do português, verbo unir:

(39)	P_1	uno	$[u'nu]$
	P_2	unes	$[u'nis]$
	P_3	une	$[u'nɛ]$
	P_4	unimos	$[ũ'nĩns]$
	P_5	unis	$[ũ'niys]$
	P_6	unem	$[ũ'nɛy]$

No espanhol, o que se tem para o mesmo verbo é:

(40)	P ₁	uno	[ˈuno]
	P ₂	unes	[ˈunes]
	P ₃	une	[ˈune]
	P ₄	unimos	[uˈnimos]
	P ₆	unen	[ˈunen]

Ao identificar os formativos no caso do espanhol, coloca-se apenas um problema, referente à vogal temática (que ocorre como i em P₄ e como e em P₁, P₂, P₃, P₆). A análise proposta por Harris é, no caso, a seguinte:

The stem is un; the person-number endings are the same as in the first and second conjugations, but the theme vowel alternates between stressed i and unstressed e. A few exceptions aside, no Spanish word has an unstressed high vowel in the final syllable (footnote: The apparent exceptions include the clitics su, tu, mi, and casi (if the last is in fact a clitic). The real exceptions include a few words of Greek origin such as énfasis, dosis, the three Latin words espíritu, tribú, ímpetu, affective words such as mami, papi, and a few miscellaneous others such as fuchi (used to indicate that something is nasty or foul smelling). Therefore we may account for the theme vowel alternation by postulating underlying thematic i and the very general rule:

$$\begin{bmatrix} V & i \\ -stress \end{bmatrix} \longrightarrow \begin{bmatrix} -high \end{bmatrix} / \text{---} C. \#$$

(Harris, 1969 : 68)

Examinemos a seguir as formas do português. O radical é identificado sem problemas como [un -], aplicando-se posteriormente à vogal a regra (4) - nasalização. Quanto à vogal temática, há alternância nas manifestações fonéticas entre [i] (P₂, P₃, P₅), [ĩ] (P₄) e [ẽ] (P₆). Já sabemos que a nasalização nestes casos é resultante da aplicação de (4) à vogal que pre

cede o segmento nasal. Resta então identificar a vogal temática, a partir das variantes [e] e [i]. Note-se desde já que em P_1 a vogal temática não se manifesta foneticamente (da mesma forma que em amar e comer), por efeito de (5) truncamento. Na variante estudada a VT ocorre como [i] em P_2 , P_3 , P_4 e P_5 e como [e] em P_6 apenas. Identifica-se |-i-| como vogal temática dos verbos da 3a. conjugação. É essa a vogal que ocorre sob acento (e. g. em P_4 e P_5). É necessário, admitindo-se que |-i-| seja a VT característica da 3a. conjugação, dar conta da variante [e] que ocorre em P_6 . Examinemos inicialmente o comportamento da vogal temática em P_2 e P_3 . Embora ela ocorra foneticamente como [i] e seja a representação de um |-i-| subjacente, há razões para suspeitar que tenha sido e em um certo ponto da derivação. Isto porque, também no português, como no espanhol, há um estágio em que se pode dizer que, a não ser por algumas exceções (e. g. iuri, jiu-jitsu, tribo (|tribut#| - cf. tribunal, tribuno, tributo, etc.), ônibus, Vênus), não há vogais finais não acentuadas que sejam altas. A ortografia reflete, inclusive, este estágio da derivação das formas, que é o que antecede a regra (7) - levantamento, responsável então pelas vogais foneticamente altas. A regra que dá conta da passagem de i a e como consequência da restrição a que fizemos menção acima, pode ser formulada como:

(41) ABAIXAMENTO

$$\left[\begin{array}{c} V \\ -\text{acento} \end{array} \right] \longrightarrow \left[-\text{alta} \right] / \text{---} C_0 \#$$

Como uma das regras mais tardias na cadeia de derivações

aplica-se então a regra (7) - levantamento, que é responsável pela ocorrência de [i] em P_2 e P_3 .

Admitida esta hipótese explica-se automaticamente a variante [e] que se manifesta em P_6 . Esta forma não apresenta condições para a aplicação de (7), porque a vogal é seguida de n, segmento não previsto no ambiente já identificado como condicionador do levantamento de vogais em posição átona final.

Identifica-se pois para o português, o tema [un+i+](radical [un-] e VT [-i-]).

As marcas número-pessoais são as mesmas já identificadas para a 1a. e a 2a. conjugações. Também o padrão acentual é o mesmo já especificado em (36).

Observa-se portanto que há novamente coincidência absoluta nas representações fonológicas básicas postuladas para o espanhol e o português.

Este aspecto, que vimos sempre procurando ressaltar, ficará ainda mais evidente quando apresentarmos em 3.4.2, a derivação das formas para ambas as línguas a partir das mesmas representações subjacentes.

Há ainda uma observação que achamos oportuno fazer neste ponto, observação esta que diz respeito às bases verbais para as três conjugações. É conveniente partir do princípio de que há três classes de bases verbais, as chamadas conjugações, caracterizadas pela vogal temática. Já fizemos, aliás, referência a estas três classes quando da introdução e consideração dos diversos dados. A noção de conjugações verbais está incorporada em todas as descrições tradicionais do sistema ver

bal da língua, não constituindo, portanto, nenhuma novidade teórica. O que queremos ressaltar aqui, entretanto, é o fato de que as vogais temáticas características de cada uma das classes de temas ou bases verbais são previsíveis e introduzidas por uma regra da mesma natureza que (35) (que dá conta da configuração fonológica dos formativos gramaticais número-pessoais).⁴¹ As bases verbais devem vir enlistadas no léxico e, como os demais formativos, são caracterizadas por um conjunto de propriedades gramático-semânticas.⁴² Uma destas propriedades é a indicação da classe a que pertencem.⁴³ Como há 3 classes, duas propriedades são suficientes para que se façam todas as distinções necessárias. Consideremos, por exemplo, os radicais |am-|, |kom -| e |un-| e as propriedades [1 conj] e [2 conj]. Tais radicais viriam indicados no léxico como se segue:

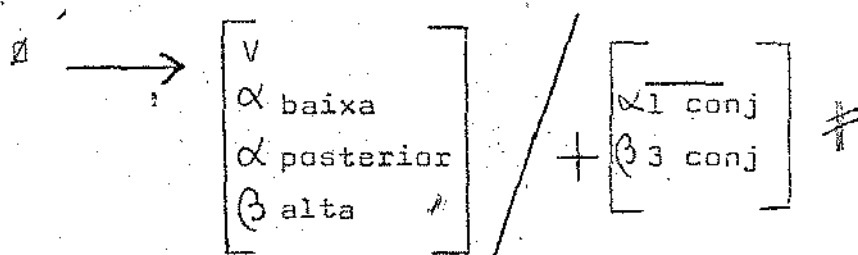
$$\begin{array}{l}
 (42) \quad |am + \begin{bmatrix} 1 \text{ conj} \\ \vdots \end{bmatrix} \#| \\
 |kow + \begin{bmatrix} - 1 \text{ conj} \\ - 3 \text{ conj} \\ \vdots \end{bmatrix} \#| \\
 |un + \begin{bmatrix} - 1 \text{ conj} \\ + 3 \text{ conj} \\ \vdots \end{bmatrix} \#|
 \end{array}$$

Observe-se que a especificação [- 3 conj] é redundante no caso dos verbos que pertencem à 1.ª classe ou conjugação (uma vez que a especificação "+" para [1 conj] os distingue já dos demais).

$$\begin{array}{l}
 (43) \quad |am +| \quad |kom +| \quad |un +| \\
 \begin{array}{ccc}
 1 \text{ conj} + & - & - \\
 3 \text{ conj} (-) & - & +
 \end{array}
 \end{array}$$

O motivo pelo qual dissemos que as vogais temáticas são previsíveis e, portanto, introduzidas por regra, é que sempre ocorrem estas três (a, e, i) das sete possibilidades existentes (cf. matriz dos segmentos vocálicos subjacentes, capítulo 1). O fato de que os outros 4 segmentos não ocorrem nunca como vogais características de temas verbais não pode ser mero acaso (como também não é mero acaso que a, e e i caracterizem certos temas nominais). Não há certeza sobre a maneira de formular a regra que fornece a configuração fonológica da VT a partir das especificações presentes no léxico, mas o formato é possivelmente semelhante a:

(4A) VOGAL TEMÁTICA



Embora seja certo que as vogais temáticas que podem ocorrer como características das classes verbais são previsíveis por regra, é importante frisar mais uma vez que não é previsível qual das três vogais ocorre com os vários radicais. Por isso há necessidade dessa indicação, no léxico, sob a forma de propriedades gramático-semânticas. É, aliás, de informações não previsíveis por regras que se constitui o léxico de qualquer língua.

Na seção seguinte aplicaremos as regras introduzidas neste capítulo para a derivação de todas as formas do presente do indicativo nas três conjugações, em espanhol e em português. Per

tiremos sempre de uma representação subjacente idêntica e indicaremos o estágio na derivação em que se obtêm as manifestações fonéticas ocorrentes no espanhol, continuando a seguir com a aplicação de mais regras para a obtenção das representações fonéticas do português.

3.4.2 Derivação das formas

Antes de apresentarmos a derivação das formas verbais consideradas neste estudo, enumeraremos as regras na ordem que deve ser seguida para a aplicação às diversas formas. Os números à direita referem-se à ordem de apresentação das regras no trabalho e, a partir deste ponto, não serão mais levadas em consideração. A ordenação aqui proposta é consequência da argumentação desenvolvida e pode, eventualmente, vir a ser modificada se o confronto de maior número de dados assim o exigir.

- (45)
- (1) VOGAL TEMÁTICA (44)
 - (2) FORMATIVOS NÚMERO-PESSOAIS (35)
 - (3) TRUNCAMENTO (5)
 - (4) ACENTO (36)
 - (5) ABAIXAMENTO (41)
 - (6) APÓCOPE (29)
 - (7) NASALIZAÇÃO (4)
 - (8) QUEDA DE n (22)
 - (9) QUEDA DE d (13)
 - (10) LEVANTAMENTO (7)
 - (11) FORMAÇÃO DE SEMIVOGAL (14)
 - (12) INSERÇÃO DE SEMIVOGAL (31)
 - (13) NASALIZAÇÃO SECUNDÁRIA (25)
 - (14) PALATALIZAÇÃO (15)

3.4.2.1 Presente do Indicativo: 1.^a conjugação regular

(46)

	P ₁	P ₂	P ₃
Regras	$ am+[+1conj]+[1^{pers}]_{-pl} \#$	$ am+[+1conj]+[2^{pers}]_{-pl} \#$	$ am+[+1conj]+[3^{pers}]_{-pl} \#$
(1)	a	a	a
(2)	o	s	—
(3)	∅	—	—
(4)	á	á	á
Espanhol	[ˈamo]	[ˈamas]	[ˈama]
(7)	õ	õ	õ
(8)	—	—	—
(9)	—	—	—
(10)	u	—	—
(11)	—	—	—
(12)	—	—	—
(13)	—	—	—
Português	[ˈõmu]	[ˈõmas]	[ˈõma]

	P ₄	P ₅	P ₆
Regras	$am + [+1conj] + \begin{bmatrix} 1press \\ +pl \end{bmatrix} \#$	$am + [+1conj] + \begin{bmatrix} 2press \\ +pl \end{bmatrix} \#$	$am + [+1conj] + \begin{bmatrix} 3press \\ +pl \end{bmatrix} \#$
(1)	a	a	a
(2)	mos	des	n
(3)	_____	_____	_____
(4)	á	á	á
Espanhol	[a'mamos]	_____ ⁴⁴	[a'man]
(7)	õ õ	õ	õ õ
(8)	_____	_____	ø
(9)	_____ \	ø	_____
(10)	u	i	_____
(11)	_____	y	_____
(12)	_____	_____	w
(13)	_____	_____	ũ
Português	[õ'mõmɐs]	[õ'mays]	[õ'mõũ]

3.4.2.2 Presente do Indicativo: 2.^a conjugação regular

(47)

	P ₁	P ₂	P ₃
Regras	$\ Kom+[-1conj]+[-3conj]+[-1pers]+[-pl]\ $	$\ Kom+[-1conj]+[-3conj]+[-2pers]+[-pl]\ $	$\ Kom+[-1conj]+[-3conj]+[-3pers]+[-pl]\ $
(1)	e	e	e
(2)	o	s	—
(3)	Ø	—	—
(4)	ó	ó	ó
Espanhol	[¹ Komo]	[¹ Komes]	[¹ Kome]
(7)	õ	õ	õ
(8)	—	—	—
(9)	—	—	—
(10)	u	i	i
(11)	—	—	—
(12)	—	—	—
(13)	—	—	—
Português	[¹ Kõmu]	[¹ Kõmis]	[¹ Kõmi]

	P ₄	P ₅	P ₆
Regras	$Kom + \begin{bmatrix} -1_{conj} \\ -3_{conj} \end{bmatrix} + \begin{bmatrix} 1_{pers} \\ +pl \end{bmatrix} \#$	$Kom + \begin{bmatrix} 1_{conj} \\ -3_{conj} \end{bmatrix} + \begin{bmatrix} 2_{pers} \\ +pl \end{bmatrix} \#$	$Kom + \begin{bmatrix} -1_{conj} \\ -3_{conj} \end{bmatrix} + \begin{bmatrix} 3_{pers} \\ +pl \end{bmatrix} \#$
(1)	e	e	e
(2)	mos	des	n
(3)	_____	_____	_____
(4)	é	é	ó
Espanhol	[Kõ'memos]	_____	[Kõ'men]
(7)	õ é	õ	õ ã
(8)	_____	_____	∅
(9)	_____ \	∅	_____
(10)	u	ı	_____
(11)	_____	y	_____
(12)	_____	_____	y
(13)	_____	_____	ỹ
Português	[Kõ'mẽmus]	[Kõ'meỹs]	[Kõ'mẽỹ]

3.4.2.3 Presente do Indicativo: 3ª conjugação regular

(48)

	P ₁	P ₂	P ₃
Regras	$un + \begin{bmatrix} -1conj \\ +3conj \end{bmatrix} + \begin{bmatrix} 1pess \\ -pl \end{bmatrix}$	$un + \begin{bmatrix} -1conj \\ +3conj \end{bmatrix} + \begin{bmatrix} 2pess \\ -pl \end{bmatrix}$	$un + \begin{bmatrix} -1conj \\ +3conj \end{bmatrix} + \begin{bmatrix} 3pess \\ -pl \end{bmatrix}$
(1)	i	i	i
(2)	o	s	—
(3)	∅	—	—
(4)	ú	ú	ú
(5)	—	e	e
Espanhol	['uno]	['unes]	['une]
(7)	ũ	ũ	ũ
(8)	—	—	—
(9)	—	—	—
(10)	u	i	i
(11)	—	—	—
(12)	—	—	—
(13)	—	—	—
Português	['ũnu]	['ũnis]	['ũni]

P₄P₅P₆

Regras	$\mu n + \begin{bmatrix} 1 \text{ conj} \\ +3 \text{ conj} \end{bmatrix} + \begin{bmatrix} 1 \text{ pess} \\ +pl \end{bmatrix} \#$	$\mu n + \begin{bmatrix} 1 \text{ conj} \\ +3 \text{ conj} \end{bmatrix} + \begin{bmatrix} 2 \text{ pess} \\ +pl \end{bmatrix} \#$	$\mu n + \begin{bmatrix} 1 \text{ conj} \\ +3 \text{ conj} \end{bmatrix} + \begin{bmatrix} 3 \text{ pess} \\ +pl \end{bmatrix} \#$
(1)	i	i	i
(2)	mos	des	n
(3)	_____	_____	_____
(4)	í	í	ú
(5)	_____	_____	e
Espanhol	[u'nimos]	_____	[u'nen]
(7)	ũ ã	ũ	ũ ã
(8)	_____	_____	Ø
(9)	_____	Ø	_____
(10)	u	i	_____
(11)	_____	y	_____
(12)	_____	_____	y
(13)	_____	_____	ỹ
Português	[ũ'nĩmus]	[ũ'nĩys]	[u'nẽỹ]

CONCLUSÃO.

O principal objetivo da exposição e análise dos dados no capítulo 3 foi o exame da hipótese segundo a qual as representações fonéticas em línguas cognatas podem explicar-se algumas vezes a partir de representações fonológicas básicas idênticas. As línguas cujas formas foram utilizadas na consideração desta hipótese foram o espanhol e o português, tendo o estudo sido limitado ao presente do indicativo nas três conjugações regulares. As conclusões a que chegamos foram as seguintes:

(a) - para a análise dos formativos gramaticais em questão podem-se formular formas básicas idênticas para as duas línguas;

(b) - as diferenças observadas no nível das manifestações fonéticas se devem ao fato de que, enquanto as formas do espanhol são obtidas em estágio relativamente próximo da derivação, é necessário prosseguir com a aplicação de mais regras para se chegar às formas do português;

(c) - é possível que, em âmbito mais geral, haja outros fatores de diferenciação superficial, tais como a ordenação diversa de um mesmo conjunto de regras, supressão de alguma (s) regra(s), etc. Não tivemos oportunidade de examinar tais possibilidades devido à limitação do corpus com o qual trabalhamos;

(d) - as mesmas conclusões ((a), (b) e (c), acima) a que chegamos com relação aos fatores de diferenciação entre duas línguas geneticamente muito próximas, são

provavelmente válidas também como medida da natureza e grau de diferenciação entre os vários dialetos de ca da uma das línguas.

A ratificação de nossa hipótese básica e das conclusões a qui apresentadas dependerá, evidentemente, da análise de um número muito maior de dados do português e do espanhol, bem como do confronto de outras línguas também em termos de relacionamento genético e estabelecimento de representações básicas. Conforme já tivemos oportunidade de mencionar por diver sas vezes, o presente trabalho é de âmbito bastante restrito. Acreditamos, entretanto, que possa ser desenvolvido oportunamente, e que novas observações interessantes com relação ao componente fonológico do espanhol e do português possam resul tar do exame de um corpus mais extenso.

Há ainda uma outra hipótese, que pode ser entendida como consequência da primeira, e com a qual gostaríamos de conclu ir o nosso estudo. É empiricamente comprovado o seguinte fato: o espanhol é mais facilmente compreendido por falantes na tivos do português, do que o português por falantes nativos do espanhol. Até o presente, tal fato tem sido mencionado como mais uma "curiosidade lingüística" sem que para ela se tenha encontrado uma explicação em termos científicos. É possí vel, contudo, que tal explicação exista. Considerando-se a hipótese segundo a qual, em alguns casos, há identidade de formas básicas e maior número de regras para o português, po- de-se dizer que:

— para os falantes nativos do português é mais fácil compre

ender certas manifestações fonéticas do espanhol porque as regras que se aplicam a estas formas fazem também parte da gramática do português. Nesse sentido, as representações fonéticas obtidas para o espanhol constituem-se em estágios intermediários na derivação das formas correspondentes do português;

- para o falante nativo do espanhol a dificuldade de compreensão das mesmas formas no português é bem maior, uma vez que algumas das regras da gramática desta língua são para eles desconhecidas, aplicando-se após já ter sido alcançado, na cadeia de derivações a partir de formas básicas idênticas, o produto final para o espanhol.

Parece-nos desnecessário ressaltar o caráter especulativo de tal hipótese. Não sugerimos que ela deva ser tomada como verdadeira a partir de tão limitado número de dados, apenas.

Achamos, entretanto, que, por ela dar conta cientificamente de um fato lingüístico até o momento não explicado, merece futura e cuidadosa consideração.

INTRODUÇÃO

1. A fonologia gerativa estuda o componente fonológico de uma gramática gerativa. Fundamenta-se basicamente na utilização de traços distintivos (noção introduzida pela Escola de Praga), para a consideração dos fenômenos ao nível da fonética e da fonêmica sistemática.

Harris (1969 : 1) indica algumas das obras que mostram o desenvolvimento desta teoria:

Chomsky, Halle e Lukoff (1956), "On Accent and Juncture in English"; Jakobson e Halle (1956), Fundamentals of Language; Chomsky (1957), Syntactic Structures; Halle (1969), The Sound Pattern of Russian; Halle (1962), "Phonology in a Generative Grammar"; Jakobson, Fant e Halle (1963), Preliminaries to Speech Analysis; Chomsky (1964), Current Issues in Linguistic Theory; Katz e Postal (1964), An Integrated Theory of Linguistic Descriptions; Chomsky (1965), Aspects of the Theory of Syntax; Chomsky e Halle (1965), "Some Controversial Questions in Phonological Theory"; Chomsky (1967), "Some General Properties of Phonological Rules"; Postal (1968), Aspects of Phonological Theory; Chomsky e Halle (1968), The Sound Pattern of English.

A última obra citada contém o mais minucioso estudo da teoria até hoje publicado, tendo servido de base a várias monografias que tratam da aplicação dos princípios da fonologia gerativa a uma língua específica.

Há ainda uma obra, publicada mais recentemente, que se ocupa de discussões metodológicas: Botha (1970), Methodological Aspects of Generative Phonology.

2. O conceito de gramática gerativa desenvolveu-se a partir dos pressupostos teóricos da teoria transformacional de análise linguística. Os princípios básicos deste modelo, elabora-

do principalmente por Noam Chomsky, encontram-se, entre outras obras, em: Chomsky (1957), Syntactic Structures e (1965), Aspects of the Theory of Syntax.

Como, para compreensão do trabalho que apresentamos, é indispensável situar com precisão o componente fonológico dentro de uma gramática gerativa, citamos aqui um trecho de Lleo (1970) que nos parece bastante claro:

A generative grammar is divided into three components: a central syntactic component which by means of phrases structure rules, transformational rules and a lexicon generates strings of formatives with their corresponding semantic features, phonological features and structural description; the semantic component and the phonological component are merely interpretive, they map those strings of formatives onto the semantic interpretation and the representation of the phonetic realization respectively.

The phonological component, therefore, assumes as given the syntax of the language and often uses its information. It consists of a partially ordered set of rules: morpheme structure rules which determine the constraints cycle, which uses the information given by the syntactic component in order to predict certain phonological properties such as stress and intonation; P-rules, which map the output of the syntactic component from its systematic phonemic representation onto its representation at the systematic phonetic level. (1)

3. Sobre a natureza da relação entre fonologia e morfologia , veja-se Schane (1968):

(...) The phonological and morphological aspects of the language are not treated as separate levels, each functioning independently of the other, but rather they are shown to be integrated into an over-all system. Morphological information is relevant for determining the phonological representation of morphemes, and conversely, phonological considerations are indispensable for establishing the morphological structure of the language. (grifo nosso). (ix-

Prefácio)

4. Convém salientar desde agora que alguns transformacionalistas já consideram discutível a necessidade de ordenação extrínseca (i.e., estabelecida com base em princípios específicos de cada língua) para as regras gramaticais. Koutsoudas, Sanders e Noll (1971) apresentam vários argumentos a favor da seguinte hipótese:

All restrictions on the relative order of application of grammatical rules are determined by universal rather than language-specific principles, from which it follows that no grammatical rules are extrinsically ordered. (1)

Admitir como verdadeira tal hipótese seria reconhecer, segundo os autores, os seguintes princípios na aplicação das regras:

- (a) - simultaneidade: — no caso dos dados satisfazerem às descrições estruturais de duas (ou mais) regras.
- (b) - ordenação: — segundo princípios universais estabelecidos a partir de características formais das próprias regras.

Uma teoria baseada em princípios universais de ordenação de regras implicaria, evidentemente, em uma maior economia para as gramáticas das línguas específicas, uma vez que tais princípios fariam parte dos chamados universais lingüísticos. Achamos, no entanto, prematuro optar por esta hipótese, tendo em vista que o problema foi posto em discussão muito recentemente e ainda não se conhecem as implicações gerais de uma possível adoção desses princípios para a teoria geral da linguagem. De qualquer forma não poderíamos nos omitir sobre o

problema, já que, se vier a ser provado que a ordenação extrínseca é realmente dispensável, certas diferenças dialetais não poderão se explicar através da ordenação diferente de regras idênticas, conforme sugerimos. Seria necessário considerar, nesse caso, a possibilidade de dialetos diferentes apresentarem pequenas diferenças na própria formulação de algumas de suas regras (dialetos mais conservadores poderiam apresentar regras menos gerais, por exemplo).

Não nos detaremos mais a considerar este problema teórico que não é, por assim dizer, essencial ao nosso trabalho. O importante é reconhecer a necessidade de uma certa ordenação nas regras (sejam extrínsecos ou intrínsecos os princípios que governam tal ordenação). Na análise que aqui apresentamos partimos do princípio de que as regras são extrinsecamente ordenadas.

5. Sabemos de apenas uma publicação em língua portuguesa que trata especificamente de princípios teóricos da fonologia generativa. Trata-se do artigo de Halle (1964) "On the Basis of Phonology", traduzido por Miriam Lemle sob o título de "Conceitos Básicos da Fonologia" em Novas Perspectivas Linguísticas, antologia de textos lingüísticos modernos publicada pela Editora Vozes (v. Lemle, 1970:115-28).

Brevemente deverá sair, pela mesma editora, a tradução de Langacker (1967-8) Language and its Structure. Neste livro, baseado em conceitos básicos da teoria transformacional, encontram-se algumas referências ao componente fonológico da gramática.

CAPÍTULO 1

1. Cf. Harms (1968):

Two levels of representation can be meaningfully distinguished: the input representations of morphemes (and of morphs), and the phonetic output — frequently called the levels of systematic phonemics and systematic phonetics respectively. No evidence for any other significant phonological level, such as taxonomic phonemic level, has been found. The systematic phonemic representation is in some ways more like the morphophonemic level of taxonomic phonemics, but cannot be identified with it. (13-4).

Algumas observações devem ser feitas a respeito do trecho acima. Harms faz referência a dois níveis significativos, apenas, que chama de nível da fonêmica sistemática e nível da fonética sistemática (nível fonológico e nível fonético, respectivamente, de acordo com a apresentação que fazemos). Diz ainda que, embora a representação fonêmica sistemática seja mais ou menos equivalente ao nível morfofonêmico da fonêmica taxonômica (modelo estrutural), não pode com ela ser identificada.

Parace-nos essencial, somente, a definição dos níveis dentro de cada teoria. As possíveis equivalências não são tão importantes assim e podem levar a conclusões falsas. O máximo que se pode ^{fazer} reconhecer que o chamado "nível fonológico" da fonologia gerativa representa um grau maior de abstração do que o "nível fonêmico" do modelo estrutural. Isso não quer dizer, no entanto, que corresponda ao "nível morfofonêmico" dessa teoria. Por este motivo preferimos, como Chomsky e Halle (1966), não fazer uso de termos como "representação morfofonêmica" ou "fonêmica sistemática", empregados com sentido diver

so em outro modelo de análise lingüística.

Neste trabalho empregamos sempre "nível fonológico" e "nível fonético", conforme definidos em 2.1.1.

Chomsky e Halle (1968) assim se manifestam sobre a inadequação de certos termos para o estudo do nível mais abstrato do componente fonológico da gramática:

Other terms that might have been used in place of the terms just proposed are "morphophonemic representation" or "systematic phonemic representation". We have avoided these terms, however, because of the technical meaning they have been given in various theories of sound structure developed in modern linguistics. The term "morphophonemic" representation seems to us appropriate only if there is another linguistically significant level of representation, intermediate in "abstractness" between lexical (phonological) and phonetic and meeting the conditions placed on "phonemic representation" in modern structural linguistics. We feel, however, that the existence of such a level has not been demonstrated and that there are strong reasons to doubt its existence. We will make no further mention of "phonemic analysis" or "phonemes" in this study and will also avoid terms such as "morphophonemic" which imply the existence of a phonemic level (11).

2. O conceito de traço distintivo vem sendo empregado nas descrições lingüísticas desde o momento em que se compreendeu não ser o fonema uma entidade indivisível, mas sim a combinação de um feixe de traços distintivos. Já é bem antiga, aliás, a noção de que os chamados sons da fala formam classes que se cruzam, classes essas identificáveis por traços comuns que determinados sons apresentam. Assim os sons nasais, por exemplo, têm em comum o fato de serem produzidos com o véu palatino abaixado, o que faz com que constituam uma classe natural, distinta das demais. O traço de nasalidade, portanto, os carac-

teriza e distingue.

O status de unidade teórica mínima só veio a ser atribuído ao traço distintivo, no entanto, com os estudos de fonologia gerativa, quando esta entidade passou a ser considerada o ponto de partida para a análise do componente fonológico.

Observe-se que, no modelo estrutural, as entidades básicas (fonemas) podem ser decompostas em traços distintivos, enquanto que no modelo gerativo as unidades básicas (traços distintivos) combinam-se em segmentos que constituem o padrão sonoro da língua. Há, pois, inversão de status entre as entidades envolvidas nestes dois modelos de descrição lingüística.

Para maiores esclarecimentos a respeito da teoria dos traços distintivos, proposta principalmente por Roman Jakobson e fundamentada em estudos do Círculo Lingüístico de Praga, ver : Jakobson e Halle (1956), Fundamentals of Language; Jakobson, Fant e Halle (1963), Preliminaries to Speech Analysis. Em Camera (1970:101-96), Fonema e Fonologia, encontra-se a tradução de um dos artigos de Jakobson e Halle, "Phonology in Relation to Phonetics", escrito em 1955 e publicado inicialmente em Fundamentals of Language. Este artigo encontra-se também em Malmberg (1970), Manual of Phonetics.

Os traços com os quais trabalhamos representam um quadro bastante modificado em relação ao que foi inicialmente idealizado por Jakobson. Baseamo-nos aqui nas modificações propostas por Chomsky e Halle (1968) e que são fruto da consideração de problemas surgidos do estudo de línguas específicas (para a definição dos traços ver principalmente o capítulo 7

de The Sound Pattern of English).

3. Os princípios teóricos que apresentamos nesta seção são discutidos com mais detalhes em Chomsky e Halle (1968 - v. principalmente partes I e IV). Se não nos restringimos aqui a indicações bibliográficas é porque achamos oportuno, por motivos expostos anteriormente, fazer acompanhar a análise de informações sobre a teoria em que ela se baseia.

4. Isto equivale a dizer que um trabalho como o que nos propomos fazer pressupõe uma análise completa do componente sintático do português e do espanhol. Tal análise, no entanto, ainda não foi feita na íntegra, tendo sido iniciados há bem pouco tempo os estudos que, dentro dessa linha teórica, incidem sobre um ou outro aspecto sintático destas duas línguas. Qualquer consideração que se faça sobre o componente fonológico do português ou do espanhol terá, por conseguinte, um caráter provisório, porque estará baseada em dados que representem o output sintático do que na maior parte das vezes se susõe serem as estruturas profundas da língua.

Esta observação nos parece imprescindível para a compreensão dos verdadeiros objetivos desse estudo, já indicados na introdução.

5. A noção de "universais lingüísticos" é consequência natural de certos postulados básicos da teoria transformacional.

Cf. Chomsky e Halle (1968):

General linguistics attempts to develop a theory of natural language as such, a system of hypotheses concerning the essential properties of any human language. These properties determine the class of possible natural languages and the

class of potential grammars for some human language. The essential properties of natural language are often referred to as "linguistic universals". Certain apparent linguistic universals may be the result merely of historical accident. For example, if only inhabitants of Tasmania survive a future war, it might be a property of all then existing languages that pitch is not used to differentiate lexical items. Accidental universals of this sort are of no importance for general linguistics, which attempts rather to characterize the range of possible human languages. The significant linguistic universals are those that must be assumed to be available to the child learning a language as an a priori, innate endowment. That there must be a rich system of a priori properties — of essential linguistic universals — is fairly obvious from the following empirical observations. Every normal child acquires an extremely intricate and abstract grammar, the properties of which are much undetermined by the available data. This takes place with great speed, under conditions that are far from ideal, and there is little significant variation among children who may differ greatly in intelligence and experience. The search for essential linguistic universals is, in effect, the study of the a priori faculté de langage that makes language acquisition possible under the given conditions of time and access to data. (4).

Quanto à natureza, podem ser divididos os universais linguísticos em "formais" e "substantivos". Cf. ainda Chomsky e Halle (1968):

It is useful to divide linguistic universals roughly into two categories. There are, first of all, certain "formal universals" that determine the structure of grammars and the form and organization of rules. In addition, there are "substantive" universals that define the set of elements that may figure in particular grammars. For example, the theory of transformational generative grammar proposes certain formal universals regarding the kinds of rules that can appear in a grammar, the kinds of structures on which they may operate, and the ordering conditions under which these rules may apply.(...) Similarly, general linguistic theory might propose, as substantive universals, that the lexi-

cal items of any language are assigned to fixed categories such as noun, verb, and adjective, and that phonetic transcriptions must make use of a particular, fixed set of phonetic features. (4).

6. Sobre as matrizes das representações fonéticas, ver Chomsky e Halle (1968):

(...) a phonetic representation has the form of a two-dimensional matrix in which the rows stand for particular phonetic features; the columns stand for consecutive segments of the utterance generated; and the entries in the matrix determine the status of each segment with respect to the features. In a full phonetic representation an entry might represent the degree of intensity with which a given feature is present in a particular segment; thus, instead of simply subdividing segments into [+strident] and [-strident], as in the example just given, the entries in the row corresponding to the feature "strident" might indicate degrees along a differentiated scale of "stridency". The phonetic symbols [p], [t], [θ], [i], [u], etc., are simply informal abbreviations for certain feature complexes; each such symbol, then, stands for a column of a matrix of the sort just described. (grifo nosso). (5).

7. Omitimos aqui as notas dos autores sobre o trecho citado, uma vez que, por fazerem referência a detalhes teóricos, não nos parecem essenciais a este capítulo, onde procuramos trazer informações sobre os aspectos teóricos básicos da teoria transformacional.

8. No léxico aparecem todos os formativos lexicais e apenas alguns formativos gramaticais. Existem formativos gramaticais que são introduzidos diretamente por regras sintáticas. Assim, não apenas o léxico, mas também as regras sintáticas, fornecem os formativos que se organizam em cadeias nas estruturas superficiais.

No capítulo 3 faremos referência a regras que introduzem certos formativos gramaticais.

9. Costuma-se também falar em forma ou representação subjacente com o mesmo sentido em que se emprega aqui o termo forma básica. A ideia é sempre a de uma representação de natureza abstrata à qual se aplicam as regras do componente fonológico.

10. Tudo parece indicar que as línguas escolhem de preferência os segmentos menos marcados, isto é, de menor complexidade. Deve haver também outros fatores que governam a escolha, mas este é ainda um campo aberto para pesquisas. Não nos estenderemos mais sobre o assunto, indicando apenas como fonte de consulta o capítulo nove de Chomsky e Halle (1968) onde o problema é abordado.

11. Não reconhecemos senão dois graus (+ ou -) para cada traço, enquanto trabalhamos no nível abstrato das representações fonológicas. Isto não quer dizer, no entanto, que no nível real das representações fonéticas não possa haver uma gradação. Na maioria das vezes, aliás, a descrição no nível fonético tem que ser feita em termos n-ários já que o que se quer representar é o grau de intensidade com que determinado traço fonético se apresenta nos segmentos (cf. nota 6, deste capítulo).

Veja-se ainda Halle (1964), que parece ter equacionado bem o problema:

A afirmação de que a atribuição de interpretações fonéticas às matrizes fonêmicas é uniforme para todas as línguas reflete o fato de que o aparelho articulatório do homem é o mesmo em toda a parte; os homens de qualquer região são capazes

de controlar os mesmos poucos aspectos do comportamento de seus aparelhos fonadores. Os traços fonéticos representam, portanto, as capacidades humanas de produzir sons vocais e constituem, neste sentido, a base fonética universal da linguagem. Como nem todos os traços fonéticos são binários (grifo nosso) o componente fonológico conterá regras que substituam alguns dos mais e menos das matrizes por números inteiros representando os diferentes graus de intensidade com que o traço em questão se manifesta na elocução. Assim, por exemplo, o fato de que o [ʌ] do inglês como em pup é menos grave ("posterior") do que o [u] de poop ficará incorporado numa regra fonológica que substitui o sinal + do traço de gravidade por um número inteiro mais alto na vogal de poop do que na de pup.

À luz do acima exposto, a extensiva discussão que tem havido a respeito da afirmativa do binarismo dos traços distintivos parece ter sido devida principalmente a uma identificação de marcas fonêmicas abstratas com os traços fonéticos a que estão associadas pelas regras de gramática. Uma vez feita a distinção entre marcas fonêmicas abstratas e traços fonéticos, há pouco campo para desacordo, pois o fato de haver mais de dois graus foneticamente distintos de gravidade não invalida a afirmação de que na representação fonêmica abstrata dos morfemas há apenas traços binários. (in: Lemle e Leite, 1970:128)

A utilização de um sistema n-ário de traços, isto é, a atribuição de valores numéricos aos traços fonéticos, está tipicamente prevista, portanto, para que se possa dar conta da capacidade do falante nativo de pronunciar a sua língua. Quando, no entanto, a pesquisa se desenvolve no nível fonêmico, e o que se procura explicar é a competência do falante quanto às generalidades de sua língua, deve-se trabalhar com um sistema binário de traços. A teoria fonológica com a qual nos propomos trabalhar não prevê a utilização de marcas fonêmicas abstratas para distinguir os segmentos opositivos de uma de -

terminada língua com base numa graduação de valores destas marcas fonêmicas abstratas. O critério do binarismo é essencial, neste nível.

12. Conforme já esclarecemos em nota anterior (v. nota 2, deste capítulo), empregamos em nossa análise os traços propostos por Chomsky e Halle (1968). No capítulo sete desta obra encontram-se definidos todos os traços. Por ser coincidente o uso que deles fazemos, achamos dispensável apresentar aqui tais definições.

Para justificar a opção que fizemos pelo emprego destes traços citamos aqui alguns trechos dos autores que nos parecem convincentes:

(a) There are several ways in which primary strictures have been treated in the phonetic literature. The most widely known approach, that of the International Phonetic Alphabet, utilizes different features to characterize the strictures in vowels and in consonants. Vowel strictures are described with the help of the features "front-back" and "high-low", whereas consonantal strictures are characterized by means of a single multivalued parameter that refers to the location of the constriction. The disadvantage of this method is that it fails to bring out the obvious parallels between vocalic and consonantal strictures. Thus, the difference between palatal and velar ^{consonants} clearly parallels that between front and back vowels, for in both cases there are the same differences in the position of the body of the tongue. There is, however, no mechanism in the IPA framework to capture this and similar facts.

One of the many contributions of R. Jakobson is a phonetic framework in which many of these parallels are properly captured. As is well known, the salient characteristic of the Jakobsonian framework is that ^{the} same three features "gravity", "compactness", and "diffuseness" — are used to describe primary strictures in both vowels and consonants. This complete iden

tification of vowel and consonant features seems in retrospect to have been too radical a solution for reasons that we briefly outline below (grifo nosso). We have therefore made a number of changes in the framework, in particular with regard to the primary cavity features. The revised framework is quite likely to appear to depart from the earlier framework much more radically than it in fact does. This deceptive impression is the result of the unfortunate need to change terminology once again and replace the by now reasonably familiar terms "compact", "diffuse", and "grave" in part by totally new terms, in part by terms that are a return to the status quo ante. We discuss the relationship between the two frameworks in section 4.2.1. (303-4)

(b) We must now inquire into the role that features "high", "low", and "back" play in the remaining class of consonants, which in terms of the present framework are anterior and/or coronal. We observe that the features may be used in a natural manner to characterize subsidiary consonantal articulations such as palatalization, velarization and pharyngealization. These subsidiary articulations consist in the superimposition of vowel-like articulations on the basic consonantal articulation. In palatalization the superimposed subsidiary articulation is [i]-like; in velarization [ɤ]-like; and in pharyngealization, [a]-like. The most straightforward procedure is, therefore, to express these superimposed vowel-like articulations with the help of the features "high", "low", and "back", which are used to characterize the same articulations when they appear in the vowels (grifo nosso). (305-6)

(c) The features discussed in the preceding sections are basically revised versions of "diffuseness", "compactness", and "gravity", which are well-known from earlier presentations of the distinctive feature framework where they served to characterize the main articulatory configurations in the vowels as well as the consonants. As more and more languages were described within this framework it became increasingly clear that there was a need for modifications along the lines discussed in the preceding section. In this section we shall examine some of the problems that arose within the earlier framework and outline the way in which these problems are overcome by the revised features presented above. This question has recently been examined also by

McCawley (1967a).

The revision proposed in the last few pages have the following main effects:

- (1) Features specifying the position of the body of the tongue are now the same for vowels and consonants.
- (2) In the characterization of vowel articulations, the features "high", "low", "back" correspond to the earlier "diffuse", "compact", and "grave" respectively. In consonants, the same three revised features correspond to palatalization, velarization and pharyngealization in the manner discussed above.
- (3) The feature "anterior" mirrors precisely the feature "diffuse" in consonants.
- (4) The feature "coronal" corresponds most closely to the feature "grave" in consonants but with opposite value. Except for the palatals ([k₁], etc.), consonants that were classified as nongrave in the earlier framework are coronal in the revised framework, whereas those that were classified as grave are noncoronal. The palatals, which in the earlier framework were nongrave, are noncoronal.

We recall that in the earlier framework the feature "diffuse" was used to characterize both the distinction between high and nonhigh vowels and that between what we have called anterior and nonanterior consonants. As a result the articulatory and acoustical characterization of the feature became quite complex and rather implausible. (See, for example, the discussion of diffuseness in Halle (1964).)

A further consequence of the same fact was the need to characterize palatalization, velarization and pharyngealization by means of independent features. This, in turn, failed to explain why these subsidiary articulations are not found with consonants that are formed with the body of the tongue, i.e., consonants that are noncoronal and nonanterior in the present framework. In the former framework this was a mere accident; in the revised framework the gap is structurally motivated, as shown in section 4.2. (grifo nosso: lembramos aqui o que foi dito, pelos autores, na seção 4.2: "The characterization of the vowels in terms of the three features above is quite straightforward and differs little from that found in most traditional phonetic books. We must observe only that the phonetic characterization of "low" and "high" rules out sounds that are [+low], for it is impossible to raise the [+high]

body of the tongue and simultaneously lower it below that level" (305)). It is worth of note that rounding (labialization), which is also a subsidiary articulation, is not subject to similar restrictions. All classes of consonants including labials, may be rounded.

A related inadequacy of the former framework is that it provided no explanation for the fact that palatalization, velarization and pharyngealization are mutually exclusive. In the revised framework the co-occurrence of these articulations is a logical impossibility since a given sound cannot be back and nonback. In the former framework, on the other hand, this is no more than a coincidence.

The former framework, furthermore, did not bring out the fact that palatalization and velarization characteristically occur before front and back vowels, respectively; the connection between palatalization and front vowels and between velarization and back vowels was no more motivated than a connection between glottalization or voicing and front vowels (grifonosso). In the revised framework, on the other hand, palatalization and velarization are obvious cases of regressive assimilation.

The earlier framework failed to account for the appearance of palatal, in place of velar, consonants in precisely the same environments where other classes of consonants undergo palatalization. (Recall that palatalization preserves the point of articulation, whereas the change of velar to palatal constitutes a change in the point of articulation). In the revised framework these two superficially distinct processes are shown to be a result of the same change, that is, [+back] to [-back]. A parallel argument can be given for the treatment of velarization and pharyngealization in the two frameworks.

The earlier framework made it impossible in principle to distinguish velar from uvular or pharyngeal consonants by means of their points of articulation. Such distinctions instead had to be made by the use of some subsidiary feature such as "stridency".

There are, however, languages (Serer, for example — see p. 305 and note 5) in which velar and uvular consonants do not differ in any such subsidiary feature and which therefore could not be accounted for. This shortcoming is easily taken care of in the revised framework, in which

the different points of articulation in velar, uvular and pharyngeal consonants are specified with the help of the feature "high", "low", and "back". (306-8).

Resta explicar o motivo pelo qual achamos que a divisão primária entre os segmentos subjacentes do português é feita através dos traços "sonorante" e "consonantal" (em vez de "vocálico" e "consonantal"). Parece-nos que as três classes assim constituídas (obstruintes, ressonantes e vogais) são mais naturais. $|m|$, $|n|$, $|l|$ e $|r|$ passam a constituir uma classe (ressonantes), o que na verdade corresponde a fenômenos observáveis na língua (e.g. o comportamento destes segmentos com relação às regras fonológicas). Se o traço distintivo fosse "vocálico" em vez de "sonorante", $|m|$ e $|n|$ estariam automaticamente separados de $|l|$ e $|r|$, já que seriam $[- \text{voc}]$ por definição, dentro do quadro em que nos baseamos, enquanto que $|l|$ e $|r|$ seriam $[+ \text{voc}]$. Cf. Chomsky e Halle (1968):

Vocalic sounds are produced with an oral cavity in which the most radical constriction does not exceed that found in the high vowels $[i]$ and $[u]$ and with vocal cords that are positioned so as to allow spontaneous voicing; in producing non-vocalic sounds one or both of these conditions are not satisfied.

Vocalic sounds, therefore, are the voiced vowels and liquids, whereas glides, nasal consonants, and obstruents, as well as voiceless vowels and liquids, are nonvocalic. (302)

Já a propriedade "sonorante" permite o agrupamento:

Sonorants are sounds produced with a vocal tract cavity configuration in which spontaneous voicing is possible; obstruents are produced with a cavity configuration that makes spontaneous voicing impossible (id. ibid., 302)

$|m|$, $|n|$, $|l|$ e $|r|$ são sons produzidos de modo a que pos

se haver vozeamento espontâneo. São, portanto, [+ sonar]. Por serem também [+ cons], constituem a classe natural dos chamados segmentos "ressonantes".

Os segmentos não-sonorantes são os chamados obstruintes. Para maiores esclarecimentos a respeito das propriedades "sonorante", "vocálico" e "consonantal", bem como para a noção da vibração espontânea ou não das cordas vocais, na qual se baseiam, inclusive, as definições de algumas destas propriedades, ver Chomsky e Halle (1968:300-3).

13. Os segmentos não-sonorantes (obstruintes) são também consonantais, mas este traço é redundante no caso destes segmentos do português. Senão, vejamos:

	<u>obstruintes</u> (consoantes verdadeiras)	<u>ressonantes</u>	<u>vogais</u>
sonorante	-	+	+
consonantal	(+)	+	-

Se os glides y e w fossem reconhecidos como segmentos subjacentes e, portanto, incluídos na matriz, seria necessário, para distinguir todos os segmentos, utilizar ainda para esta primeira divisão a propriedade "vocálico". Ter-se-ia, então:

	obstruintes	ressonantes	vogais	glides
sonorante	-	+	+	+
consonantal	(+)	+	-	-
vocálico	(-)	(-)	+	-

14. Para que seja entendida a especificação de [m] e [n] como re-

dundantemente não contínuos, lembramos que trabalhamos aqui com as propriedades conforme definidas por Chomsky e Halle (1968). Os dois autores assim definem contínuo/não contínuo:

CONTINUANT - NONCONTINUANT (STOP)

In the production of continuant sounds, the primary constriction in the vowel tract is not narrowed to the point where the air flow ^{is} blocked; in stops the air flow through the mouth (grifo nesso) is effectively blocked.

Among the stops are the plosives (nasal as well as oral), the affricates, and the glottal stops, as well as various types of sounds with closure not only at the point of primary constriction but also at supplementary constriction, including clicks, other doubly articulated plosives (labiovelars), and implosive and ejective stops. (317-8)

Com base no acima exposto $[m]$ e $[n]$ são não contínuos por definição, portanto (o que não quer dizer que, considerados com base em outras possíveis definições das propriedades, não possam ser considerados contínuos). A razão pela qual abordamos aqui o problema desses dois segmentos em particular é que tem havido muita discussão a respeito da especificação adequada das nasais com relação à continuidade. Neste trabalho mantemos a especificação proposta por Chomsky e Halle (1968).

15. Achemos que a distinção entre as obstruintes com a mesma especificação para os traços "contínuo", "coronal" e "anterior" é feita, no português, através do traço "tenso" (e não através de "vozeado", conforme já tem sido proposto em análises da língua). Chegamos a tal conclusão principalmente pelos motivos que se seguem:

(a) - em posição postônica final é realmente a tensão e não

o vozeamento que mantém a distinção em casos como [ˈgaty] (gato): [ˈgat̪y] (gado) onde [y] representa uma realização surda da vogal em posição final átona e [t̪] uma realização não tensa de consoante nesta mesma posição. Neste ambiente o traço "vozeado", redundante para os segmentos [-tensos], não se manifesta foneticamente.

Com base na observação deste fato, alguns autores classificam já as consoantes da língua em fortes e lenes (o que equivale, em termos da nossa análise, a tensas e não tensas).

Cf. Lamle (1963):

É interessante assinalar, ainda, que esses mesmos ambientes condicionam variação alofônica também na classe das consoantes lenes: /b/, /d/, /g/, /v/, /z/, /ʒ/, /m/, /n/, /ɲ/, /l/, /ʎ/, /r/, /y/, /w/, manifestam-se com alofones surdos quando precedem os alofones surdos das vogais, e com alofones sonoros quando em ambiente sonoro. Com isto não se neutraliza, como poderia parecer, o contraste entre /b/ e /p/, /c/ e /t/, /g/ e /k/ ou entre as duas séries de fricativas, nestas posições, pois a diferença entre fortis e lenis continua a separá-las.

Tal verificação traz implícita uma conclusão importante para a apreensão dos traços distintivos deste sistema fonológico, pois mostra que não é a oposição surda-sonora que é relevante, e sim a fortis-lenis, o que constitui uma quebra da tradição gramatical corrente. (33)

Cf. ainda Pontes (1965):

Classificamos as consoantes em fortes e lenes e não em surdas e sonoras, como outros autores costumam fazer, em vista da ocorrência de alofones surdos das consoantes lenes (v.2.4.). Em casos como o de /ˈtaba/ [ˈtap̪A] frente a /ˈtapa/ [ˈtapA], desaparece a oposição surdez-sonoridade, permanecendo, para distinguir as consoantes, a oposição de leng-forte. Assinalamos que as consoantes fortes são sempre surdas e as lenes apresentam alofones sonoros exceto em ambien

te surdo, descrito em 2.4. (27);

(b) - Partindo-se da hipótese de que o traço é distintivo para as consoantes, é possível postular regras mais gerais para dar conta de certos detalhes fonéticos. Pode-se dizer, por exemplo, que os segmentos não tensos manifestam o traço [-voz] em posição átona final. Uma regra assim geral não seria descoberta caso se considerasse distintiva a sonoridade em vez da tensão para as consoantes do português.

16. A especificação dos traços redundantes pode também ser feita por meio de regras. Nesse caso, em vez das indicações "+" ou "-" entre parênteses, apareceriam "zeros" (0) nas matrizes e os traços redundantes seriam fornecidos por regras. Se tivesse sido esta a representação escolhida para (1) as regras seriam apresentadas discursivamente como:

- (a) - (i) - Os segmentos não sonorantes são consonantais e não nasais;
- (ii) - Os segmentos não sonorantes, não contínuos e coronais são anteriores;
- (iii) - Os segmentos não sonorantes contínuos não coronais são anteriores;
- (b) - (i) - Os segmentos sonorantes consonantais são anteriores e não tensos;
- (ii) - Os segmentos sonorantes consonantais nasais são não contínuos;
- (iii) - Os segmentos sonorantes consonantais não nasais são coronais.

Poderiam ainda ser formalizadas da seguinte maneira:

$$(a) - (i) - \begin{bmatrix} - \text{son} \end{bmatrix} \longrightarrow \begin{bmatrix} + \text{cons} \\ - \text{nas} \end{bmatrix}$$

$$(ii) - \begin{bmatrix} - \text{son} \\ - \text{cont} \\ + \text{cor} \end{bmatrix} \longrightarrow \begin{bmatrix} + \text{ant} \end{bmatrix}$$

$$(iii) - \begin{bmatrix} - \text{son} \\ + \text{cont} \\ - \text{cor} \end{bmatrix} \longrightarrow \begin{bmatrix} + \text{ant} \end{bmatrix}$$

$$(b) - (i) - \begin{bmatrix} + \text{son} \\ + \text{cons} \end{bmatrix} \longrightarrow \begin{bmatrix} + \text{ant} \\ - \text{tenso} \end{bmatrix}$$

$$(ii) - \begin{bmatrix} + \text{son} \\ + \text{cons} \\ + \text{nas} \end{bmatrix} \longrightarrow \begin{bmatrix} - \text{cont} \end{bmatrix}$$

$$(iii) - \begin{bmatrix} + \text{son} \\ + \text{cons} \\ - \text{nas} \end{bmatrix} \longrightarrow \begin{bmatrix} - \text{cor} \end{bmatrix}$$

Conforme se pode observar, haveria três regras para dar conta dos traços redundantes dos segmentos obstruintes (i.e., não sonorantes), ou seja: (a) - (i, ii, iii). Por intermédio de alguns recursos abreviatórios poder-se-ia chegar a uma simplificação quanto ao número de regras, fundindo inicialmente (ii) e (iii) em:

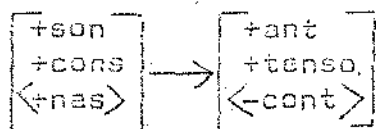
$$(ii)' : \begin{bmatrix} - \text{son} \\ \alpha \text{cont} \\ - \alpha \text{cor} \end{bmatrix} \longrightarrow \begin{bmatrix} + \text{ant} \end{bmatrix}$$

Em seguida, considerando conjuntamente (i) e (ii)', o resultado seria uma única regra para os traços redundantes das obstruintes:

$$(a) - \begin{bmatrix} - \text{son} \\ \alpha \text{cont} \\ - \alpha \text{cor} \end{bmatrix} \longrightarrow \begin{bmatrix} + \text{cons} \\ - \text{nas} \\ + \text{ant} \end{bmatrix}$$

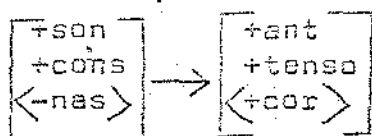
Quanto aos segmentos sonorantes, as três regras (b) - (i, ii, iii) poderiam ser abreviadas em:

(i)' (fusão de (i) em (ii)):

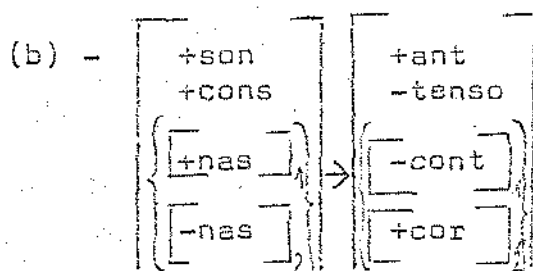


e:

(i)'' (fusão de (i) e (iii)):



(i)' e (i)'' poderiam ainda se resumir em (b), que seria então a única regra necessária para dar conta dos traços redundantes das sonorantes:



(Isto é, se o segmento sonorante consonantal for nasal, será redundantemente não contínuo, além de anterior e não tenso. Se for não nasal, a propriedade redundante, além das duas já mencionadas, será coronal).

Se mostramos em nota estes exemplos de formalização de regras é porque temos assim oportunidade de apresentar, nesta etapa introdutória do trabalho, certas informações sobre recursos abreviatórios cuja utilização poderá se mostrar necessária no decorrer da análise.

Observe-se, por exemplo, que no momento de formalizar as regras, as subdivisões (ii) e (iii) de (a) abreviaram-se em uma só regra:

$$(ii): \begin{bmatrix} -\text{son} \\ \alpha\text{cont} \\ -\alpha\text{cor} \end{bmatrix} \longrightarrow \begin{bmatrix} +\text{ant} \end{bmatrix}$$

Isto foi possível com a utilização da variável α (alfa) cujo emprego na formalização de regras em fonologia gerativa procuraremos explicar a seguir.

Considerem-se inicialmente as regras:

$$\begin{array}{l} (1) \begin{bmatrix} +F_1 \\ -F_2 \\ +F_3 \\ +F_4 \end{bmatrix} \longrightarrow \begin{bmatrix} +F_5 \end{bmatrix} / \text{ x } \text{---} \text{ y} \\ (2) \begin{bmatrix} +F_1 \\ -F_2 \\ -F_3 \\ +F_4 \end{bmatrix} \longrightarrow \begin{bmatrix} -F_5 \end{bmatrix} / \text{ x } \text{---} \text{ y} \end{array}$$

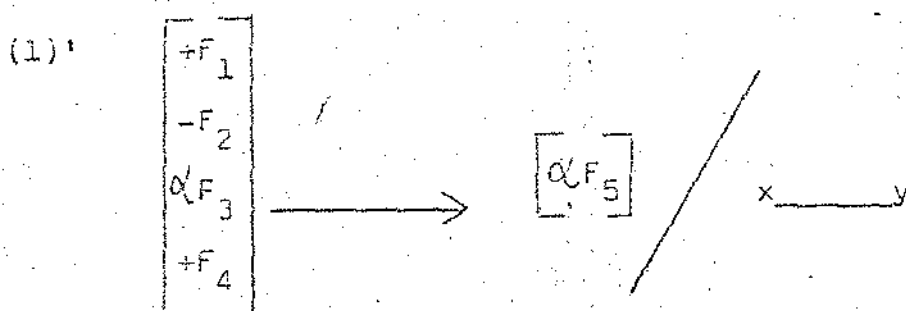
F significa propriedade (ingl. "feature") e os índices representam cada uma das propriedades em questão. ".../x---y" é a indicação de um ambiente qualquer (onde a barra horizontal indica o próprio segmento especificado à esquerda da seta e as letras X e Y, segmentos ou cadeias de segmentos que o precedem ou seguem. Podem também ser interpretadas como nenhum segmento).

Nota-se logo, ao observar conjuntamente (1) e (2), que há uma correlação entre $[F_3]$ e $[F_5]$, pois a única diferença entre as duas regras é com relação a essas duas propriedades (i.e., onde existe + em (1), à esquerda da seta, existe - em (2)). Embora os valores sejam diferentes, no caso, as propriedades envolvidas são exatamente as mesmas em (1) e (2). É justamente

com base nesta simetria que se pode fazer uma generalização importante (correlação entre propriedades) e, para expressar tal generalização, utiliza-se a letra grega α como uma variável ($\begin{bmatrix} + \\ - \end{bmatrix} F$), a ser interpretada segundo o critério seguinte:

- se o valor atribuído a α for + em uma parte da regra, ela terá que representar sempre + em qualquer outra parte da regra onde apareça;
- se o valor atribuído a α for -, será sempre -.

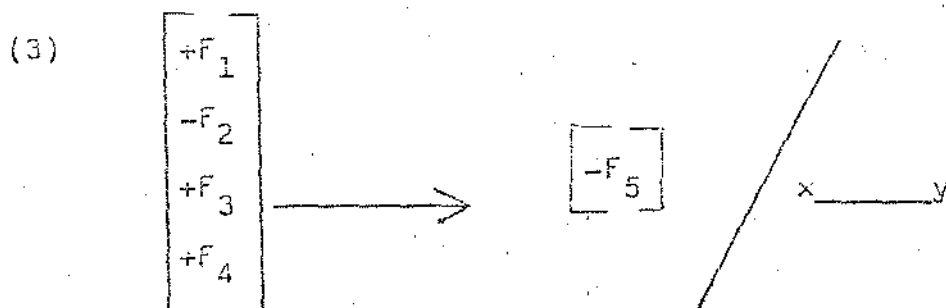
Utilizando a variável, temos então a abreviação de (1) e (2) em:



(i.e., sempre que o segmento com as demais especificações à esquerda for $\begin{bmatrix} +F_3 \end{bmatrix}$, será também $\begin{bmatrix} +F_5 \end{bmatrix}$. Sendo $\begin{bmatrix} -F_3 \end{bmatrix}$, será $\begin{bmatrix} -F_5 \end{bmatrix}$).

Quando queremos indicar, em uma parte da regra, que o valor de α é o oposto daquele que lhe foi atribuído inicialmente, utilizamos a notação $\underline{-\alpha}$.

Assim, se em vez das situações descritas em (1) e (2) tivéssemos:

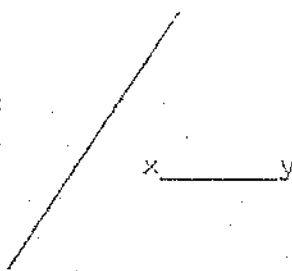


(4)

$$\begin{bmatrix} +F_1 \\ -F_2 \\ -F_3 \\ +F_4 \end{bmatrix}$$



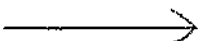
$$\begin{bmatrix} +F_5 \end{bmatrix}$$



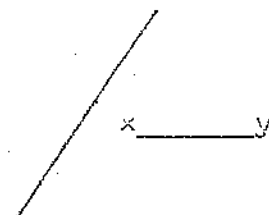
a abreviação possível seria:

(3)

$$\begin{bmatrix} +F_1 \\ -F_2 \\ +F_3 \\ +F_4 \end{bmatrix}$$



$$\begin{bmatrix} -F_5 \end{bmatrix}$$

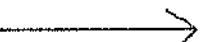


(onde se indica que a correlação entre as propriedades é em termos de $[+F_3] \rightarrow [-F_5]$ ou $[-F_3] \rightarrow [+F_5]$).

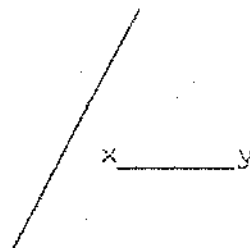
Pode existir também correlação entre um número maior de propriedades. Assim:

(5)

$$\begin{bmatrix} +F_1 \\ -F_2 \\ +F_3 \\ +F_4 \end{bmatrix}$$

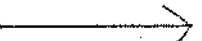


$$\begin{bmatrix} +F_5 \end{bmatrix}$$

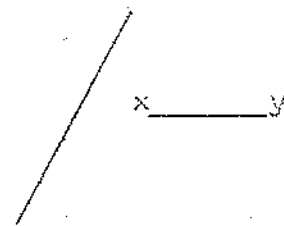


(6)

$$\begin{bmatrix} +F_1 \\ -F_2 \\ -F_3 \\ -F_4 \end{bmatrix}$$



$$\begin{bmatrix} -F_5 \end{bmatrix}$$



De onde temos:

(5)'

$$\begin{bmatrix} +F_1 \\ -F_2 \\ \alpha F_3 \\ \alpha F_4 \end{bmatrix} \longrightarrow$$

$$\boxed{\alpha F_5}$$

x — y

183

Observe-se que a correlação entre $\begin{bmatrix} F_3 \end{bmatrix}$ e $\begin{bmatrix} F_5 \end{bmatrix}$ e entre $\begin{bmatrix} F_4 \end{bmatrix}$ e $\begin{bmatrix} F_5 \end{bmatrix}$ é a mesma, nas duas regras (i.e. $\begin{bmatrix} +F_3 \\ +F_4 \end{bmatrix} \longrightarrow \begin{bmatrix} +F_5 \end{bmatrix}$, e $\begin{bmatrix} -F_3 \\ -F_4 \end{bmatrix} \longrightarrow \begin{bmatrix} -F_5 \end{bmatrix}$). Se a correlação existisse em outros termos, seria necessário utilizar $-\alpha$ à esquerda, para uma das propriedades:

As regras:

(7)

$$\begin{bmatrix} +F_1 \\ -F_2 \\ +F_3 \\ -F_4 \end{bmatrix} \longrightarrow$$

$$\boxed{+F_5}$$

x — y

(8)

$$\begin{bmatrix} +F_1 \\ -F_2 \\ -F_3 \\ +F_4 \end{bmatrix} \longrightarrow$$

$$\boxed{-F_5}$$

x — y

abreviar-se-iam então em:

(7)'

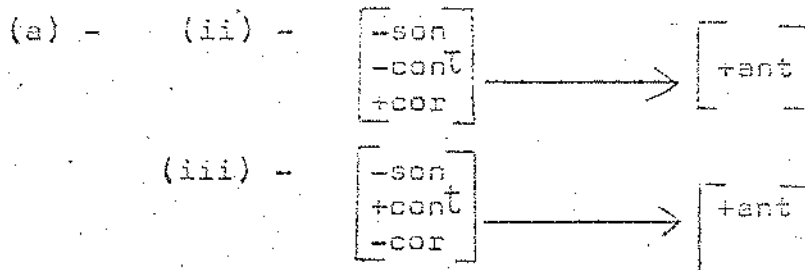
$$\begin{bmatrix} +F_1 \\ -F_2 \\ \alpha F_3 \\ -\alpha F_4 \end{bmatrix} \longrightarrow$$

$$\boxed{\alpha F_5}$$

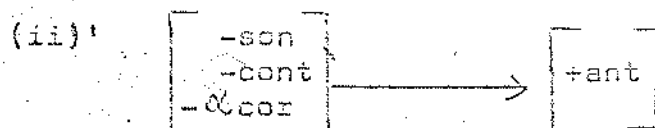
x — y

(i.e., com relação a $\begin{bmatrix} F_3 \end{bmatrix} \longrightarrow \begin{bmatrix} F_5 \end{bmatrix}$ onde $\begin{bmatrix} +... \end{bmatrix} \longrightarrow \begin{bmatrix} +... \end{bmatrix}$ e $\begin{bmatrix} -... \end{bmatrix} \longrightarrow \begin{bmatrix} -... \end{bmatrix}$; empregamos α ; no caso de $\begin{bmatrix} F_4 \end{bmatrix} \longrightarrow \begin{bmatrix} F_5 \end{bmatrix}$ onde temos $\begin{bmatrix} +... \end{bmatrix} \longrightarrow \begin{bmatrix} -... \end{bmatrix}$ e $\begin{bmatrix} -... \end{bmatrix} \longrightarrow \begin{bmatrix} +... \end{bmatrix}$, necessitamos de $-\alpha$).

Explica-se assim o que nos permitiu abreviar.



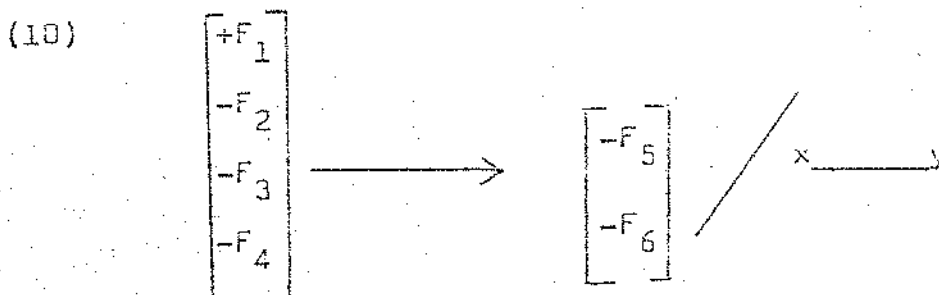
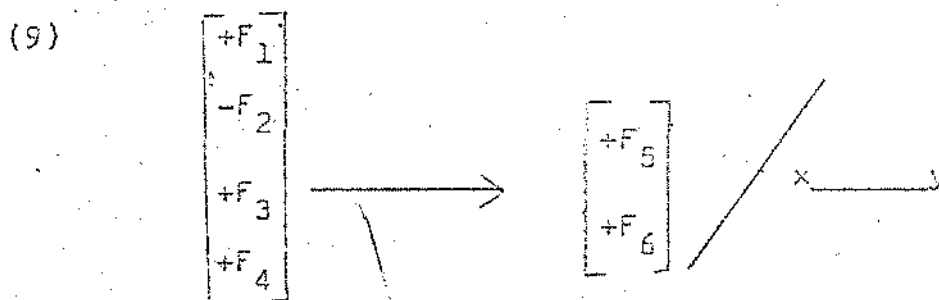
em:



A situação é paralela à das regras (7) e (8), só que a correlação existe apenas entre duas propriedades à esquerda. À direita, o valor é sempre o mesmo.

Há ainda outro caso interessante que merece ser considerado. Trata-se da independência entre as correlações, que só pode ser expressa por meio de variáveis diferentes (α, β, γ , etc.)

Considerem-se, por exemplo:



(11)

$$\begin{bmatrix} +F_1 \\ -F_2 \\ +F_3 \\ -F_4 \end{bmatrix} \longrightarrow \begin{bmatrix} +F_5 \\ -F_6 \end{bmatrix} \quad \text{---} \quad x \text{ ---} y$$

(12)

$$\begin{bmatrix} +F_1 \\ -F_2 \\ -F_3 \\ +F_4 \end{bmatrix} \longrightarrow \begin{bmatrix} -F_5 \\ +F_6 \end{bmatrix} \quad \text{---} \quad x \text{ ---} y$$

Se existissem apenas as regras (9) e (10), a abreviação se resolveria com a utilização de α , apenas. No entanto, considerando as quatro regras em conjunto, vê-se que as correlações existentes são entre $[F_3]$ e $[F_5]$ e entre $[F_4]$ e $[F_6]$. Através das variáveis α e β , reunimos em uma só regra as quatro combinações possíveis:

(9)

$$\begin{bmatrix} +F_1 \\ -F_2 \\ \alpha F_3 \\ \beta F_4 \end{bmatrix} \longrightarrow \begin{bmatrix} \alpha F_5 \\ \beta F_6 \end{bmatrix} \quad \text{---} \quad x \text{ ---} y$$

Por mostrar de modo claro as correlações entre certas propriedades, tal tipo de notação, através do emprego de letras gregas como variáveis, permite que fiquem patenteadas generalizações importantes relativas a processos fonológicos operantes na língua.

O uso da notação " $\langle \dots \rangle$ ", por nós utilizada na regra (9), é explicado por Harms (1968):

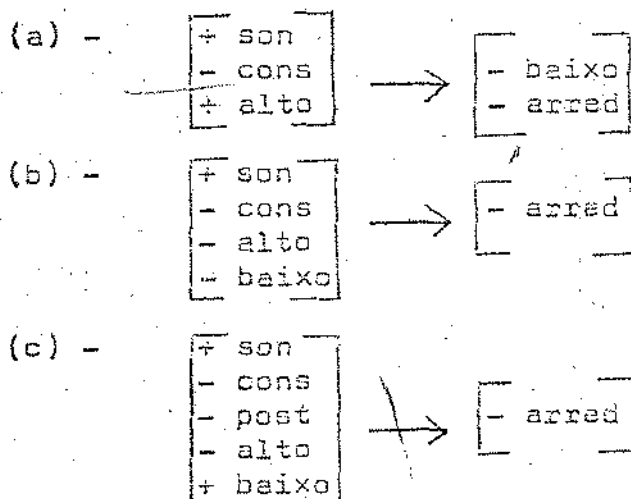
The 'angled parentheses' $\langle \dots \rangle$ indicate elements in one part of the rule that require the

presence of elements within angles in another part of the rule (56).

Deduz-se daí que, em (a), os segmentos não consonantes são consonantais e não nasais e, além disso, se forem ^{Também} $\langle \alpha \text{ cont} \rangle$ $\langle \alpha \text{ cor} \rangle$ serão anteriores.

Existam ainda várias convenções utilizadas para reduzir regras semelhantes a uma só, ou mesmo para suprimir informações por vezes irrelevantes. A esse respeito, ver o esquema apresentado por Harms (1968 : capítulo 7, "Abbreviatory devices").

17. Caso os traços redundantes tivessem sido indicados por "zero" (0) na matriz das vogais, as regras para suprir estes traços seriam:



Podem-se agrupar as três regras através de recursos abreviatórios mas a formulação pareceria bem mais complicada, além de ser desnecessária para os nossos propósitos.

18. Cf. Halle (1964) a respeito da utilização dos traços nas matrizes e das modificações que se podem operar nestas representações através das regras do componente fonológico:

No que expusemos acima, os traços distintivos foram utilizados para dois diferentes propósitos.

Por um lado, serviram para caracterizar vários aspectos do comportamento do aparelho fonador, tal como a localização dos diversos estreitamentos, a presença ou ausência de vibração das cordas vocais, o levantamento ou abaixamento do véu palatino, e assim por diante. Por outro lado, os traços funcionaram como marcas abstratas para a designação de morfemas individuais. A este ponto torna-se necessário explicar como essa dupla função dos traços se enquadra na teoria.

Já se disse que as regras que constituem o componente fonológico de uma gramática relacionam uma matriz que consiste de marcas abstratas — a representação fonêmica — com uma matriz onde cada marca representa um aspecto particular do comportamento do aparelho fonador. Esta última é o nosso equivalente da transcrição fonética convencional. Na representação fonêmica os diferentes traços só podem assumir dois valores, positivos ou negativos. No entanto, nenhum conteúdo fonético está associado diretamente aos traços, que aí funcionam como marcadores diferenciais abstratos.

As regras do componente fonológico modificam — por vezes bem radicalmente — as matrizes da representação fonológica: as regras atribuem valor a traços não-fonêmicos, mudam o valor de certos traços, e suam uma interpretação fonética as fileiras individuais da matriz (grifo nosso). A interpretação fonética dada as fileiras da matriz é uniforme para todas as línguas: i.e., alguma fileira na matriz estará associada ao traço vocálico — não vocálico, outra ao traço consonantal — não consonantal e assim por diante. Este fato explica a nossa prática de designarmos as fileiras nas matrizes fonêmicas que representam marcas diferenciais abstratas pelos nomes dos diversos traços fonéticos. Quando designamos uma dada fileira na matriz fonêmica pelo nome de um dado traço fonético, estamos implicitamente antecipando que a gramática irá acabar por associar essa fileira com o traço fonético em questão (127-8).

(v. também nota 11, deste capítulo).

CAPÍTULO 2

1. Indicam-se apenas as diferenças formais características de cada grupo. Assim, embora se observe que em hospital (I) e hospede (II) ao /t/ da primeira forma corresponde o /d/ da segunda, a correspondência que queremos ressaltar entre as formas do grupo H é entre i e e, e por isso apenas esta diferença é assinalada neste grupo.

2. A formulação de algumas dessas regras será apresentada oportunamente, no capítulo seguinte, à medida que os dados forem sugerindo os processos fonológicos envolvidos. No presente estágio da exposição é suficiente constatar a existência de tais regras.

3. Para uma discussão detalhada de como a teoria transformacional procede à categorização dos formativos, no léxico, com relação à estratificação vocábular, ver Chomsky e Halle (1968: 172-7, 373-80).

Argumentos a favor do reconhecimento de estratos lexicais são apresentados também por Harris (1969 : 127-3) Schane (1968: 26-30) e Postal (1968 : 129-39), entre outros. Em Postal encontra-se a mais completa discussão teórica do assunto.

4. Esta afirmativa pode parecer estranha, à primeira vista, uma vez que a ênfase é dada sempre ao estudo das formas do estrato II, principalmente em trabalhos diacrônicos. Isto impede, por assim dizer, que se constate o status minoritário deste conjunto de formativos. No entanto, uma pesquisa em qualquer dicionário, após o conhecimento das regras básicas da fonologia do português, levará fatosamente ao reconhecimento dos

to fato, perfeitamente explicável, aliás (v. a este respeito, 2.4).

5. Observe-se que as propriedades discriminativas podem também ser introduzidas através de regras de reajuste (ingl. 'readjustment rules' - cf. Chomsky e Halle, 1968 : 371-2, 376).
6. Não discutiremos aqui a teoria de marca e as convenções universais de interpretação [uf] e [mf] (onde u significa não marcado, ingl. 'unmarked', m significa marcado, ingl. 'marked' e f significa uma propriedade qualquer, ingl. 'feature'). A esse respeito ver o capítulo 9 de Chomsky e Halle (1968), onde é abordado o assunto do conteúdo intrínseco das propriedades, e onde são apresentadas algumas das convenções universais para interpretação das marcas.

Achamos oportuno, no momento, observar apenas que:

- (a) - as marcas são também utilizadas para indicar, no léxico, a filiação ou não dos formativos a determinadas categorias e subcategorias. Assim, [mf] e [uf], no caso dos estratos que vimos considerando no português;
 - (b) - u e m podem, ao se aplicar uma convenção interpretativa para determinada propriedade, receber a especificação [+f] ou [-f]. Isto é, o fato de uma forma ser não marcada ([uf]) com relação a determinada propriedade não quer dizer necessariamente que ela será, por convenção, [-f]. Poderá ser, conforme o caso, [+f], assim como uma forma marcada poderá ser interpretada como [-f]. As especificações adequadas dependem, portanto, das regras interpretativas. Fazemos aqui esta

Observação porque é justamente neste aspecto que a teoria da marca proposta pelos transformacionalistas difere da teoria da marca conforme entendida por Troubetzkoy (1970 : 77) (onde o elemento não marcado é interpretado sempre como $[-F]$ e o marcado como $[+F]$). No caso específico da propriedade discríptica $[E]$ o que acontece é que a interpretação das marcas vem a ser, coincidentemente:

$$\begin{array}{ccc} [mE] & \longrightarrow & [+E] \\ [uE] & \longrightarrow & [-E]; \end{array}$$

- (c) - no léxico, apenas os formativos marcados com relação a determinada categoria ($[m...]$) recebem uma especificação concreta (que, como vimos, poderá ser + ou -, dependendo das convenções de interpretação das marcas). Os elementos não especificados são automaticamente interpretados como não marcados ($[u...]$) para essa categoria.

7. A notação $[E]$ ('special'), por nós utilizada na análise do português, corresponde a $[5]$ ('special') nas citações que aqui fazemos de textos em inglês.

8. Exemplos concretos dos casos mencionados nesta seção serão vistos oportunamente, à medida que forem sendo introduzidas algumas das regras básicas da fonologia do português. Ficará então comprovado que, somente com base na estratificação do léxico, pode-se postular um conjunto de regras que reflitam a simplicidade e generalidade dos processos fonológicos.

O emprego das notações $[-\bar{a}]$ e $[\bar{a}]$ se explicará melhor quando forem considerados casos específicos de regras que apresentam $[\bar{a}]$ na descrição estrutural.

9. Schane (1968) faz indicações explícitas a respeito dos itens lexicais que podem formar tais substratos (que são, como vimos, subclasses morfológicas) nas línguas:

(...) Foreign words, proper names, dialectal loans, etc., will need to be noted as such in the lexicon as they invariably do not obey the same phonological constraints as the mass of words which form the native stock. Since particular phonological rules will not necessarily be applicable to all morphemes in the language, morphemes must be assigned to different classes. This is analogous to the situation in morphology where verb stems, for example, are assigned to different classes for the purpose of conjugation. (128).

(...) The distinction between learned and nonlearned is precisely the same type of vocabulary division. Such divisions are arbitrary to the extent that the phonological shape of forms does not in any way determine the particular phonological processes undergone by them; hence, the forms need to be marked in some other way. (26).

10. Harris (1969) baseia-se em princípios psicolinguísticos ao argumentar a utilização de informação diacrônica como ponto de partida para análises sincrônicas:

(...) Language learners do not have access to historical data as such. Consequently, such data can in no way figure in the motivation or justification of a synchronic grammar. (grifo nosso). ^{Based} It is entirely another matter when grammars wholly on syntactic data happen to reflect, to whatever extent, historical processes. What has motivated the use of the feature $[S]$ in the present study is simply the fact that certain clear synchronic generalizations will be lost if a relatively small set of forms is not distinguished that undergo a relatively small set of rules not applicable to other forms. (129).

11. Foley (1965, capítulo 1) baseia-se também em dois estratos lexicais ("erudito" e "vulgar") para o estudo da morfologia do espanhol. A distinção, segundo ele, é tradicionalmente reconhecida e se faz necessária em um estudo sincrônico da língua. O autor não comenta, no entanto, a escolha dos termos empregados na categorização dos estratos.

CAPÍTULO 3

1. Citam-se, entre os aspectos teóricos que mais têm suscitado discussões nos últimos tempos, o problema do grau de abstração das formas básicas postuladas e a possibilidade de utilização ou não de conceitos como "neutralização absoluta" (i. e., quando um segmento subjacente postulado não se manifesta nunca foneticamente. Ver, a esse respeito, o artigo de Paul Kiparsky, 1968, "How Abstract is Phonology"?). Também o problema da existência ou não de ordenação estrínseca das regras fonológicas tem sido motivo de grande controvérsia. Nota-se que as discussões baseiam-se sempre em problemas surgidos do estudo de línguas específicas. (Cf. Kiparsky, op cit.; Hyman, 1970; Kisseberth, 1969; Brame, 1969; Anderson, 1969, 1970; Koutsoudás, Sanders e Noll, 1971).
2. Entre os trabalhos de orientação estruturalista destacam-se os seguintes:
Camara, 1969, 1970, 1972; Eastlack, 1969; Pontes, 1965.
3. Ainda não há nenhum trabalho publicado que trate especificamente de aspectos fonológicos e morfológicos da classe verbal no português, dentro de uma abordagem transformacional.
4. Os chamados formativos gramaticais podem vir enlistados no léxico ou ser introduzidos através de regras sintáticas, conhecidas como "spell-out rules". Algumas dessas regras serão aqui apresentadas, oportunamente.
5. Distinções de voz não são, portanto, relevantes para este estudo, já que não podem ser atribuídas a nenhum formativo gramatical. Observa-se que os rótulos utilizados no diagrama in

trouduzido em 3.2.1.5 referem-se a várias noções gramaticais (número, pessoa, tempo, modo, aspecto), mas não à noção de voz.

A falha da NGE ao não distinguir noções gramaticais operantes em diferentes níveis da língua não escapou despercebida à maioria dos gramáticos que, forçados por decreto oficial a incluir a voz entre as flexões verbais do português, apressam-se a indicar o seu conhecimento de que a chamada "flexão de voz" é na verdade um processo sintático na língua (Cf. Carmo et al. 1971 (vol. II) : 47, nota 3).

6. Cf. a) O povo parece não acreditar mais em políticos.

b) Estes óculos custaram uma fortuna.

7. Um estudo mais detalhado das noções gramaticais de tempo e aspecto foge aos propósitos deste estudo. Partimos aqui de categorias aspecto-temporais que podem, em linhas gerais, ser consideradas tradicionais. Justificamos o emprego de tais categorias como ponto de partida apenas, já que a nossa principal preocupação não é aqui a análise da sintaxe verbal. Modificações podem vir a ser necessárias como resultado de estudos mais específicos nesse campo. Euhl (1960) e Rallides (1971) representam discussões interessantes sobre as noções de tempo e aspecto.

8. A respeito do futuro em português, v. Camara (1970). Aspectos diacrônicos e sincrônicos do problema são apresentados nesta obra, de cunho estruturalista.

9. O rótulo anterior foi aqui introduzido para distinguir entre as duas formas marcadas como [+ perf] em português (antes, anteriormente). Caracteriza-se assim o tradicional "pretérito mais-

que-perfeito", que representa um evento anterior a outro evento também passado. Como já observamos anteriormente, tais fatos requerem motivação sintática independente, da qual, por razões também já expostas, não nos ocuparemos neste trabalho.

10. Mencionamos aqui as obras indicadas por Morris (1969) para estudo mais detalhado da sintaxe verbal do espanhol: Lenz (1944 : pp.424-473), Bull (1960) e Stockwell, Bowen e Martin (1965 : pp.132-165).

Trabalhos de orientação transformacional dedicados exclusivamente à sintaxe verbal do português inexistem até o momento. O assunto, no entanto, é abordado em vários estudos sobre diversos aspectos do componente sintático da língua surgidos nos últimos anos.

11. Observe-se que Camara não leva em conta a noção gramatical de aspecto ao apresentar a fórmula para o vocábulo verbal no português. Como já observamos anteriormente (nota 7 deste capítulo) não abordaremos neste trabalho o problema da verdadeira significação gramatical dos vários sufixos. Lembremos apenas que parece realmente não ser possível segmentar um sufixo aspectual na língua, e que este noção talvez seja expressa cumulativamente através do sufixo modo-temporal (que poderia então ser considerado aspecto-modo-temporal).

12. Achamos interessante citar aqui um trecho de Langacker (1972) onde o autor explicita o conceito de motivação inerente em uma análise lingüística:

When dealing with a restricted range of data, a linguist can often conceive of several alternative ways, all fairly plausible, of account -

ing for the facts. What is needed in such a situation is some way to avoid having to choose arbitrarily among alternative analyses. Presumably, only one analysis is correct, and the linguist would like to have a principled basis for determining which one. (grifo nosso).

Because a language is a tightly integrated system, rules and representations postulated to account for one range of data frequently turn out to be useful in the description of other data that previously seemed unrelated to it. A rule or underlying representation that has this property is said to be independently motivated. The independent motivation of rules and representations is a powerful criterion for choosing among alternative analyses. Faced with a choice between two analyses that account equally well for a given body of data, a linguist will not hesitate to choose the one that involves rules and underlying representations that are independently motivated by virtue of helping to account for other data (grifo nosso). By considering the consequences of competing analyses for the rest of grammar he is constructing, the linguist is thus able to choose among them on a principled basis.

The independent motivation of rules and underlying representations is an extremely important component of linguistic argumentation. To justify a proposed set of rules and representations, it is not sufficient merely to show that they work, that they account for the data — in most instances, there are other analyses that will also work. Finding an analysis that works is of limited interest unless there is some reason to believe that the analysis makes true claims about the language. There are grounds for believing this when the proposed rules and underlying representations are shown to be independently motivated, for this means that evidence from different facets of the language converges to favor one particular analysis. If an analysis formulated to account for one group of facts were incorrect, there would be no reason to expect it to be of any value in expressing the regularities implicit in other data (27-28).

13. O critério de naturalidade deve governar qualquer análise gramatical. Na verdade o que se busca não é um conjunto de soluções arbitrárias, mas sim a sistematização de como a lin-

que organiza de fato os dados. Veja-se a esse respeito o que diz Klima (1959) em artigo no qual aborda o problema da negação em inglês:

In speaking of grammatical elements, categories, and relationships emerging in a natural way, I mean to specify a nonarbitrariness that is a concomitant both of the objective of grammar rules and of the criterion of economy as adopted above; that is, that symbols representing grammatical categories are postulated only if the presence of such symbols accounts for some ultimate difference in the shape of some other combination of symbols (i.e., with another analysis) not including the symbol in question. Thus, all forms are considered the same until some formal reason to differentiate among them presents itself, and once in possession of a grammatically differentiated feature by necessity in one usage, a form will be assumed to have that feature everywhere else as well, pending the discovery of formal reasons to differentiate further. What cannot be known to be the same nonetheless can — and, in accordance with the principle of economy, must — be assumed the same until a difference is demonstrated. (...) nothing is relevant except what occurs in a natural way in the grammar as a whole. It is continuous and exclusive reference to the total system that prevents the analysis of a segment, isolated on a more or less notional basis and thus arbitrary (as far as the grammar is concerned), from being merely the more or less rigorous description of categories from traditional grammar. On the other hand, if the grammar itself contains distinctions analogous to those categories, then the existence of such analog is at least one measure of how interesting and revealing the analysis is. (248-9)

14. Por estarmos utilizando a análise proposta por Harris para o espanhol como termo de comparação, deixamos de levar em conta as formas da 2a. pessoa do plural, já que o autor não se ocupa deste complexo número-pessoal. Gostaríamos de mencionar mais uma vez, no entanto, que a possibilidade existe, de tais formas ocorrerem (ainda que marginalmente), e elas devem, portanto, ser descritas.

15. O português apresenta também certos ditongos e tritongos orais em posição final, que não se manifestam nas demais línguas românicas. (Cf. Saciuk, 1970).

16. A utilização na análise de regras como estas representa uma maior vantagem do ponto-de-vista descritivo:

(...) an enormous descriptive advantage is already apparent over the analysis of inflections into strings of morphemes: features are independently available for mention either as the input or as the environment of a rule, and phonologically irrelevant features need not be mentioned at all (grifo nosso). It is possible to capture a vast number of generalizations when the full set of such rules is considered, generalizations that are unstatable in terms of ordered strings of morphemes. (Harris, 1969:67)

Observe-se que as regras dão conta apenas das formas fonologicamente marcadas para número-pessoa, não havendo portanto necessidade de uma regra para o complexo P3, fonologicamente representado por Ø.

17. É importante observar desde já que dois tipos de nasalização resultam no português, uma considerada fonêmica e outra fonética, não distintiva. As vogais nasais consideradas fonêmicas são a manifestação na superfície de seqüências de vogal seguida do segmento nasal n na mesma sílaba, nas representações fonológicas básicas. (Notamos aqui que a noção de estrutura silábica pode ser relevante para a consideração de determinados processos fonológicos como queda de segmentos, assimilação, etc. Tal possibilidade não foi ainda considerada com atenção em estudos de fonologia gerativa. Uma hipótese, por exemplo, é a da existência de regras diacrônicas e mesmo síncronicas de mudança de estrutura silábica. Não nos dete-

recomos, no momento, a examinar o assunto). O segmento nasal, após causar a nasalização da vogal precedente, cai ou assimila o ponto de articulação da consoante seguinte. Explicam-se assim contrastes existentes na língua como tampa - [tãpa] - [tãpa] : tapa - [tapa] - [tãpa]; mão - [mãõ] - [mãõ] : mau - [mau] - (a queda do l se dá provavelmente por efeito de uma regra como a que faz cair o n) - [maw]. As vogais que apresentam nasalização fonética são as que adquirem o traço nasal pela regra que nasaliza as vogais diante de qualquer segmento nasal, e que é de aplicação geral na língua.

Já Mattoso Camara (1970) chamava a atenção para o problema:

A língua portuguesa se caracteriza, entre as línguas românicas, por uma emissão nasal das vogais muitas vezes. O mesmo fato se apresenta em francês; mas em condições fonológicas um tanto diversas, como veremos mais tarde. Nas demais línguas românicas, o que a fonética registra é uma leve nasalização de uma vogal em contacto com uma consoante nasal da sílaba seguinte, no mesmo vocábulo.

Já vimos que nessa posição as vogais portuguesas, no Brasil, sofrem uma redução, mesmo do seu quadro tônico, com [ã] abafado (levemente posterior) e a neutralização das oposições /ê/ : /ë/ e /ô/ : /ö/ em proveito das médias de segundo grau. Vemos agora que, como nas demais línguas românicas, também podem ficar nasalizadas.

Mas não há equivalência entre as duas emissões nasais. O segundo tipo de nasalidade não funciona para distinguir formas, e não é, portanto, de natureza fonológica. (...)

A ressalva tem muita importância, porque o português, ao lado da nasalidade fonológica, também pode ter essa nasalidade, ocorrente por assimilação à consoante nasal de uma sílaba seguinte. É preciso assinalar, portanto, que uma nasalidade como de junta, oposta a luta, ou

de cinto, oposto a cito, ou de landa, oposto a loda, e assim por diante, não se deve confundir com uma pronúncia levemente nasal da primeira vogal de ano, ou de cino, ou de una, ou de tano, etc., em que o falante tende a antecipar o abajamento do véu palatino, necessário à emissão da consoante na sílaba seguinte, e emite já nasalada a vogal precedente. Ai, não há oposição entre a vogal nasalada e a vogal, também possível, sem qualquer nasalização. Com a nasalização, ou sem ela, aparecerão sempre as mesmas formas vocabulares, ano, cino, una, tano, etc. (grifo nosso). (...) Diante de uma possível nasalização, que é meramente mecânica e fonética (sem efeito para distinguir formas da língua) e uma nasalização que se opõe distintivamente à não nasalização, é preciso encontrar um traço específico que caracterize as vogais que são nasais em termos fonêmicos. São elas as únicas vogais nasais portuguesas que merecem tal classificação.

O meu ponto de vista, já antigo (Camara 1953, 59 ss.), que ainda não foi aceito pacificamente, é que se deve procurar esse traço distintivo na constituição de sílaba. Em outros termos: a vogal nasal fica entendida como um grupo de dois fonemas, que se combinam na sílaba — vogal e elemento nasal. (36-37).

Dentro do esquema teórico no qual vimos trabalhando, é interessante sustentar o ponto-de-vista de Camara de que as vogais nasais fonêmicas resultem da sequência Vn na mesma sílaba. Observe-se que a sua análise, embora feita em moldes estruturalistas, traz implícita, no caso das vogais nasais, a consideração de uma forma subjacente a partir da qual se explicam as manifestações superficiais.

A regra geral de nasalização de vogal antes de segmento nasal já foi por nós apresentada (4). A regra posterior de queda de n ou assimilação de ponto de articulação será introduzida quando considerarmos o problema das ditongas nasais do português (análise da marca número-pessoal em P_6).

13. Existe uma outra possibilidade de análise no caso das re-

ses nominais, i.e., pode-se considerar [assin-] como motivo lexical ao qual se acrescentam os formativos gramaticais [-o] e [-a] do gênero masculino e feminino, respectivamente. Nesse caso, não havendo a base nominal [assine-] por ocasião do acréscimo da marca de feminino, a regra do truncamento não se aplicaria, evidentemente. Existam, supondo, alguns dados sincrônicos que sustentem, de certa forma, a hipótese de que a noção do gênero masculino está associada ao [-o] (e.g. a existência de variantes como princípio por origin-ine, a criação de formas masculinas como corruço, barbeirão, etc., para os chamados substantivos epicenos). Não nos parece interessante, no entanto, adotar a análise acima mencionada para dar conta desses dados, uma vez que, se não se necessesse a existência de bases nominais terminadas em vogal, restaria por explicar um número considerável de itens terminados em e e que são masculinos ou femininos. A análise através da qual se reconhecem bases nominais caracterizadas por vogais específicas (como no caso das três conjugações verbais) implica em uma maior simplificação do corpus, portanto, pelas razões que se seguem:

— evidencia-se o paralelismo entre bases nominais e bases verbais (caracterizadas por determinadas vogais), as quais se acrescentam formativos gramaticais de gênero e número (no caso dos nomes) e de aspecto, modo, tempo, número e pessoa (no caso dos verbos);

— Considera-se apenas uma marca de gênero ([+a], para o feminino), que pode ser acrescentada a algumas das bases co-

minais. Em termos formais é bastante, portanto, reconhecer a existência de um formativo gramatical de gênero feminino. A distinção semântica entre os dois gêneros é sempre feita, no entanto, indicando-se o gênero da determinação formativo lexical por ocasião da sua especificação em termos de propriedades gramático-semânticas (Cf. Klima, 1989 : 312, observações sobre GSF - "grammatico-semantic features" : "The ultimate lexical items carry along with them not only a string of phonemes but also a bundle of such features").

Um formativo lexical como menino, por exemplo, teria, entre outras, especificações como as seguintes:

+ Nome
+ Humano
+ Animado
+ Masculino

19. Há evidência sincrônica para que se considere trissilábico ou dissilábico o infinitivo desses verbos. Pode-se dizer, por exemplo, que o infinitivo de ler tem como representação sistêmica $|\text{le}^{\text{h}}\text{e}^{\text{h}}\text{r}^{\text{h}}|$ porque existe na língua, juntamente com as formas deste verbo, o adjetivo legível, cujo conteúdo semântico é sem dúvida relacionado a ler. Não nos deteremos aqui a fazer especulações em torno das possíveis representações fonológicas para todos esses verbos porque isto estaria fora do âmbito do presente trabalho. Citamos aqui o exemplo de ler para que se tenha uma idéia do tipo de considerações que co-

vem ser feitas quando se quer estabelecer representações subjacentes.

20. Observa-se que na primeira hipótese indicamos $[-is]$, e não $[-ys]$ como forma básica. Nesse caso, a semivogal que ocorre nas manifestações fonéticas seria explicada por uma regra geral do português que inverte a vocalidade de um dos segmentos marcados como $[+voc]$ na sequência VV . Esta regra seria formulada logo a seguir porque é também aplicável na derivação a partir de $[-des]$. O que queremos enfatizar aqui é o motivo porque não seria escolhida a representação fonológica com a semivogal ($[-ys]$), na primeira hipótese, e o motivo principal é que a ocorrência das semivogais é sempre previsível na língua, não sendo necessário postular como subjacentes os segmentos $[y]$ e $[w]$.

21. A razão pela qual propomos $[-des]$ e não $[-dis]$ como representação sistematizada é que, no caso de credes, ledes, vades, etc., onde o d é mantido, há variantes da língua onde nas manifestações fonéticas ocorre e e não i. Nesse caso é mais econômico manter o e nas representações subjacentes a postular uma regra de levantamento dessa vogal em posição final não cantuada, aplicável a certos dialetos, apenas.

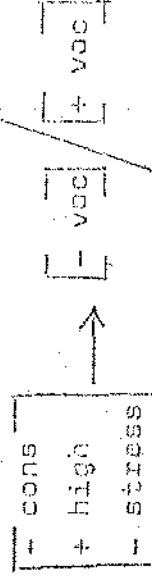
22. A regra apresentada aqui como queda de d é, na verdade, mais geral, havendo exemplos também de queda de y nas mesmas circunstâncias. (Cf. lex-lexível). Pode ser que haja alguma sincronia para se dizer ainda que outros segmentos estão sujeitos à aplicação de mesma regra (e. g. y). Não examinamos aqui o problema, uma vez que não está diretamente relacionado.

não ao assunto que vimos discutindo. A regra que formulamos (10), preferimos não especificar em termos de traços distintivos, deixando assim implícita a hipótese de que outros segmentos podem vir a ser incluídos na regra. Costâncias ainda de observar que, por esse mesmo motivo (possibilidade de generalização das regras pelo confronto de maior número de dados), indicamos na maioria das regras apresentadas neste trabalho, apenas os segmentos envolvidos, sem uma explicação completa em termos de traços distintivos. A especificação de cada um é óbvia; entretanto, a partir das matrizes apresentadas no capítulo I.

23. Saciuk (1970), também se ocupa da regra de formação de glides no português:

This GLIDE FORMATION rule applies to unstressed high vowels, preceded or followed by a vowel, and converts them into the corresponding glides:

(...) GLIDE FORMATION



The vowels to which this rule applies may be oral or nasal, as we saw in the derivation of [paʃs] (...). Again, the neighborhood convention is being used in this rule, which implies that the order of the two environments in which the rule applies would have to be first /V___ and then /___V. This would seem to be correct for forms like guidado [kujdaðu] 'caro', suicida [sujsidal] 'suicidal', etc., but there are other forms which would suggest that the ordering of the two environments has to be reversed, e. g. ciunar [sjunar], 'to be jealous', the ending -rio [-rju], etc. Suicida itself may also appear phonetically as [suisidal]. In the vast majority of the cases the ordering of the two environments is not crucial, since one of the two contiguous vowels will be either stressed or non-

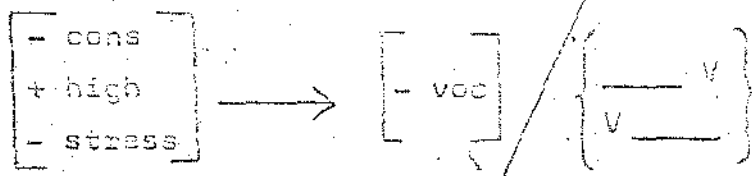
high in which case the rule will apply only to the unstressed high vowel /e, g. bol [boɫ] 'ox', sear [seər] 'to spin'). But, when at the time of application of this rule there are two consecutive unstressed high vowels, the ordering of the two environments becomes very important. At this time, it is not clear to me if they should be ordered as originally proposed by Sachs or in the opposite way. It might be interesting to point out that in Spanish the GLIDE FORMATION rule will have to be stated in such a way that it will apply first in the environment / V and then / V (see Sadiuk 1969 : 77-8). Culdad and ciudad, for example, are realized phonetically as [kwɪdad] and [sjudad] in Spanish.

Não há dúvida de que, nos exemplos citados por Sadiuk onde há uma sequência de duas vogais especificadas como [reata], ambas não acentuadas, é necessário que haja uma converção para que a regra se aplique à vogal que realmente se manifesta foneticamente como semivogal. A regra de formação de vogais no português, conforme proposta por Sadiuk e por nós apressentada neste trabalho, deve ser interpretada levando-se em conta a "convenção de vizinhança" (ingl. 'neighborhood convention') no ambiente. Esta converção, proposta por Sachs (1966), diz que quando uma regra fonológica é formulada sem a letra indicadora de ambiente precedente ou seguinte a um certo segmento, ele pode se aplicar em ambas as circunstâncias, mas na seguinte ordem:

- X → Y / Z
1. X → Y / Z
2. X Y / Z

Note-se que a observação feita por Sadiuk sobre os exemplos do espanhol leva a crer que naqueles casos a ordem na aplicação aos dois possíveis ambientes é exatamente a inversa

do que prevê a convenção de vizinhança. Harris (1969:127) assim apresenta a versão final da regra da formação de glides para o espanhol:



Pode ser que no português o ambiente também deva ser interpretado nesta ordem, em alguns casos (e. g. em suicida, citado por Sadiuk, que pode ocorrer foneticamente como [swi'sida]).

24. Há razões sincrônicas para postular a representação sistêmica [am+a+re+des#] para P₅ no futuro do Subjuntivo e infinitivo flexionado. Apesar de estar entre vogais na representação subjacente, o d é mantido na representação fonética porque, ao se aplicar a regra que faz cair esta segmento entre vogais nos formativos [+E] (lembramos aqui que todos os formativos gramaticais são considerados [+E]), o d que precede o e na representação fonológica desta forma já caiu por efeito de uma regra de apócope, generalização que será por nós discutida oportunamente ((29)). O d não se encontra mais em ambiente intervocálico, portanto, a (13) não se aplica.

25. Para o emprego da propriedade alta na caracterização do efeito de palatalização em alguns dos segmentos obstruintes, cf. Chomsky (1968 : 304-305). Cf. também, no presente trabalho, nota 12 c: - capítulo 1.

26. Pães, mãos, irmãos, leões e capitães manifestam-se foneticamente como [ˈpãys], [ˈmãws], [iɪˈmãws], [ˈlɔys] e [kapitãys], respectivamente. Como vimos (nota 20 deste capítulo

1a) a ocorrência das semivogais é sempre previsível, na português, não havendo razão para que figurem nas representações fonológicas. Assim sendo partimos de representações como [pãns+s#], [mãns+s#], [irmãns+s#], [leons+s#] e [kapitãns+s#] e, após a aplicação das regras de nasalização (4) e queda do segmento nasal (ainda não discutida), aplica-se à vogal final a regra do levantamento (7) e, finalmente, a de formação de semivogal (14).

27. Não mencionamos até o momento a regra de acento por causa da ordem que vimos seguindo na apresentação do material. Note-se desde já que tal regra deve preceder quase todas as apresentadas até agora. Quando tratamos, por exemplo, do problema do levantamento de vogais em posição final, mencionamos o fato de que as vogais às quais tal regra se aplica são não acentuadas. Assume-se que a regra de acento já tenha sido aplicada às formas porque, como se vê, ela fornece ambiente para (7). Ocupar-nos-emos do problema de acentuação imediatamente após a discussão da marca número-pessoal em P_S .

28. O i não está presente na representação fonológica desta forma (como também em veia - [venɛ#]), sendo posteriormente inserido por regra fonológica. Este é, com efeito, um processo regular na língua. Uma semivogal com a mesma especificação de posterioridade que a vogal precedente é inserida sempre que há uma sequência de duas vogais, a primeira das quais é acentuada (e. g. [a'reya], [i'veja], [pa'seyu], [i'riɣu], [i'voɣu], etc.). Na variante em que nos baseamos o fenômeno é bem geral. Não podemos deixar de observar, entretanto, que em outras va-

riantes do português a regra só se aplica em ambiente de vogais não-altas (e. g. [a'reya]). Em uma forma como frio, após a aplicação da regra de levantamento (7), aplica-se logo (13) - formação de semivogal e o que se tem é [ʔfriu]. Isso parece indicar que, no caso destas variantes, a regra de inserção de semivogais (que apenas mencionamos nesta nota, não tendo chegado a uma formulação) deve ser ordenada após (13). Este vem a ser então um exemplo de como certas variantes regionais se explicam pela ordenação diferente das mesmas regras.

29. É importante observar que as formas de (21) I b. representam já uma etapa subsequente, sendo o resultado de uma regra mais tardia de desnasalização de algumas das vogais nasalizadas por efeito de (4). Não procuraremos estabelecer aqui os ambientes onde ocorre tal desnasalização, pois tal fenômeno não é relevante para a explicação das formas verbais propriamente ditas, i.e., da maneira como se combinam os diversos formativos lexicais e gramaticais nos vários grupos de formas que constituem os paradigmas.

Para uma discussão mais detalhada dos ambientes onde possivelmente ocorre tal desnasalização, v. Saciuk (1970:209 e ss.).

30. Observe-se que o segmento nasal cai, não apenas entre vogais, mas também entre vogais e segmentos contínuos como f, s, [∨]s, [∨]v, [∨]z, [∨]ʒ, r, (e. g. ênfase [ʔẽfazi]; anzol [ʔʒ'zow], etc). Mantém-se, porém, quando está entre uma vogal e um segmento não contínuo. Assimila, nesse caso, o ponto de articu-

lação deste segmento (e. g. campo ['kâmpu], canto ['kântu], canguru ['kângu'ru], etc.). Uma regra das mais tardias dá conta da assimilação, por n, do ponto de articulação do segmento seguinte. (Esta é, possivelmente, uma das chamadas regras fonológicas universais, aplicável à maioria das línguas).

Através do uso da propriedade [contínuo] no ambiente de (22), agrupam-se os ambientes em que ocorra a queda, uma vez que todas as vogais são redundantemente contínuas.

31. Para a justificação de o e e nas representações fonológicas básicas destas formas, v. nota 27 deste capítulo.

32. Saciuk (1970) assim explica a regra de nasalização secundária no português:

This rule nasalizes vowels (and glides) when they are preceded or followed by a nasalized vowel.

(...) SECONDARY NASALIZATION

$$[-\text{cons}] \longrightarrow [+ \text{nasl}] / \left[\begin{array}{c} \text{V} \\ + \text{nasl} \end{array} \right]$$

The environment of this rule is given according to the 'neighborhood convention' proposed by Bach (1968). Thus, the rule applies in the environment $/ \left[\begin{array}{c} \text{V} \\ \text{---} \\ \text{V} \end{array} \right] /$. We have to specify that this nasality

assimilation takes place only in the environment of a nasal vowel and not a nasal consonant, because there are thousands of words with syllable initial nasal consonants (at the point in the derivation at which this rule applies and also in the phonetic output) that do not nasalize the following vowels, e.g. morar [morár] 'to reside', maleta [maleta] 'suitcase', nada [náda] 'nothing', etc., etc. Note that in mínimo [mínimu] 'minimal' and namorado [nãmorádu] 'in love' (masc.), for example, the vowels followed by nasal consonants are nasalized, while those preceded by nasal consonants are not. Very accurate measurements with mechanical devices indicate some nasalization in vowels preceded by N, but the degree of nasalization in this case is weaker than in the

vowels that undergo the rules of NASALIZATION, PROGRESSIVE NASALIZATION, or SECONDARY NASALIZATION. It is very probable that our theory of phonology eventually will have to talk about different degrees of nasality and not just about [+nasal] or [-nasal] vowels, i.e. this feature would become scalar instead of binary (as more and more features are being shown to be). The highest degree of nasality would appear in vowels that occur between two nasal consonants in the phonetic representation, followed by the nasalization of the vowels that undergo the three rules just mentioned. The weakest degree of nasalization would be found in vowels which are just preceded by nasal consonants in the phonetic output. (204-205).

33. A regra de acentuação para os nomes terá naturalmente que dar conta dos formativos proparoxítonos (que são em número bastante reduzidos como formativos [+E], ou pertencem a categorias identificáveis como [-E]). Os formativos oxítonos são, em sua maioria, atribuíveis a um outro estrato lexical identificável na língua, o dos empréstimos de origem indígena. Uma vez reconhecida a existência da vogal subjacente e em vários formativos foneticamente oxítonos e a regra de apócope em razão da qual tais pseudo-oxítonos existem, verifica-se que a grande maioria do léxico é realmente constituída de paroxítonos.

34. A esta vogal, em posição final não acentuada, aplica-se posteriormente a regra (7) - levantamento, responsável pelo resultado fonético [ˈklubi] e [tˈimi], respectivamente (a segunda forma aplica-se também a regra (15) - palatalização).

35. Observe-se o comportamento diferente do segmento l nos dois formativos abaixo:

a. Mal |male#| (Cf. maledicência, malevolência, etc.).

— pl. malas.

b. sol |sole#| (Cf. solar, ensolarado, etc.).

— pl. sóis.

Nas formas do plural o l cai em sóis, mantendo-se, no entanto, em malas. A explicação para este fato, neste ponto de exposição, parece-nos já bastante evidente. (a) é um formativo $[-E]$ e (b) é especificado como $[+E]$. Por ser $[-E]$, (a) não se submete a uma outra regra existente na língua, queda de l, que faz parte do estoque de regras aplicáveis a formativos $[+E]$, apenas. Quanto a esta regra de queda de l, basta para o momento reconhecer a sua existência e notar que o segmento que cai está sempre entre vogais, e nos formativos $[+E]$. Não nos deteremos aqui a tentar uma formulação adequada, mesmo porque talvez deva também ser levado em conta na formulação desta regra a noção de mudança de estrutura silábica (v., a esse respeito, a nota 7 deste capítulo).

36. A formulação que aqui apresentamos para a regra de apócope é apenas uma primeira aproximação. Deverá sem dúvida ser reformulada após a consideração de um número maior de dados, uma vez que há casos de e's que se mantêm em alguns dos ambientes indicados como condicionadores da queda. (e.g. vale (nome e verbo), pára (verbo), etc.). Pode ser que seja necessário incluir mais alguma especificação na regra para estabelecer uma distinção entre os e's que caem por efeito de (29) e os que se mantêm. Talvez esta especificação possa ser feita em termos da propriedade $[tenso]$, mas não nos ocuparemos no presente trabalho em examinar mais detalhadamente tal possibilidade.

dade.

Também não nos parece seguro estabelecer agora o âmbito de aplicação da regra de apócope (i.e., decidir se ela se aplica apenas a formativos [+E] ou a todos os formativos da língua), antes de estudar detidamente o problema, considerando inclusive um número mais significativo de dados.

37. Só há um caso em que, depois de vogais nasalizadas, em posição final, não se insere um glide homorgânico. Trata-se de um grupo de formas como rã, lã, irmã, etc. A regra de inserção não se aplica, nesses casos, simplesmente porque, por ocasião de sua aplicação, as formas apresentam uma vogal não nasalizada em posição final. Ilustraremos este ponto com as derivações seguintes:

	I	II
	pane #	lana #
(29)	pan #	—
(4)	pãn #	lãna #
(22)	pẽ #	lã a #
(31)	põw #	—
(25)	põw #	lãõ #
CRASE	—	lõ #
REDUÇÃO	—	lõ #

Até o momento de aplicação de (31) - inserção de semivocal, a vogal final da forma apresentada em II é ainda a, não nasalizada. Após a aplicação desta regra, então, observa-se que há nos dois casos ambiente para (25) nasalização secundária (no caso de II, a vogal nasalizada por efeito desta regra passa também a ser não baixa, conforme previsto na formulação). A lãõ aplicam-se então duas regras que não discutiremos aqui em

detalhe. Estas regras são:

- (a) - crase de vogais idênticas (do que resulta uma vogal foneticamente longa);
- (b) - redução (cujo output é uma vogal foneticamente breve).

Observe-se ainda que essas duas regras podem teoricamente ser reunidas em uma só, i.e., pode-se formular a regra de crase de tal forma que o output seja uma vogal breve. É possível, entretanto, que a regra de crase faça parte do conjunto de regras chamadas universais, e que haja línguas nas quais ocorra apenas a crase de vogais idênticas, mantendo-se longa a vogal resultante. Por esse motivo preferimos nos referir aqui a duas regras em vez de uma, apenas.

Explica-se, assim, a existência de certas vogais nasalizadas em posição final, no português, após as quais não se insere uma semivogal. É que, quando se obtêm tais vogais nasalizadas, a regra (31), inserção de semivogais, já passou na série de derivações.

38. O reconhecimento de que há inserção de glides também após vogais altas em posição final permite uma formulação mais geral para esta regra, i.e., exclui do ambiente a especificação [-alta]. Saciuk (1970) menciona a possibilidade, mas formula a regra de modo a que sejam inseridos glides apenas após vogais finais nasalizadas não altas. Não deixa, entretanto, de mencionar em nota a possibilidade de generalização:

At the time of application of this rule, only [ã], [ẽ], [ĩ], and [ũ] appear in this position. It seems to me that careful phonetic investigation would show that, at least in some dialects of Portuguese, word final [ũ] and [ĩ]

also acquire a homorganic glide in word final position. Thus, the phonetic representations of fim 'end', the masculine indefinite article um, and atum 'tuna' would be [fĩĩ], [ũũ], and [atũũ], respectively. If such were the case, then the GLIDE INSERTION rule could not only be simplified by omitting the feature specification [-high] in the structural description, but it would become much more general. (220-221).

39. Esta análise apresentada para os ditongos nasais do português é uma possibilidade, apenas. Não podemos deixar de reconhecer que há outras análises possíveis. Uma delas seria considerar que em posição final o segmento nasal n passa a ser não sonorante e não consonantal, isto é, ɲ. Isso tornaria a descrição mais econômica, sob certos aspectos, porque haveria, em vez de três regras (queda de nasal, inserção de glide, nasalização secundária), apenas uma (e. g. em [kom+e+nɲ] - [kõmẽɲ]) ou, em certos casos, duas (e. g. em [am+a+nɲ] - [ãmãɲ] - [ãmãũ]) quando há necessidade de fazer com que a semivogal resultante da regra que chamaremos de "n → semivogal nasalizada" concorde em posterioridade com a vogal precedente.

Por enquanto tal possibilidade é produto de mera especulação, não tendo sido examinadas com o devido cuidado as possíveis implicações de sua incorporação à gramática da língua. A chamamos oportuno mencioná-la aqui, já que pode abrir caminho para considerações futuras.

40. A regra de acentuação proposta dá conta do padrão acentual de certas formas verbais, apenas (daí a necessidade da indicação]_v, na regra). Um estudo mais completo, que incluísse todas as possibilidades de formas verbais na língua, nos -

traria, no entanto, que ela é válida para a grande maioria das formas, havendo apenas alguns casos de acento na antepe - núltima sílaba (e. g. amávamos, amáramos). Para dar conta de tais casos é bastante provável que devam ser incluídas na regra indicações de ordem morfológica (cujo uso em fonologia está se tornando cada vez mais freqüente, pois a idéia da existência de classes que constituem o domínio de certos fenômenos de natureza fonológica parece realmente ter certo fundamento. Para fins da presente discussão é suficiente que a regra seja formulada como em (36).

Também não fazemos referência aqui à regra de acentuação para os nomes, que deve ser formulada de maneira diferente, levando-se em conta inclusive outras propriedades (como, talvez, a ^{tensão} das vogais) para o estabelecimento preciso dos ambientes onde pode recair o acento.

41. Tais regras são conhecidas como "spell-out rules" porque inserem material fonológico nas representações de acordo com informações fornecidas por propriedades gramático-semânticas.

42. A respeito destas propriedades, v. nota 18, deste capítulo.

43. Harris (1969) assim se manifesta a respeito do problema das classes verbais:

The conjugational class of a verb is unpredictable, that is, it must be listed as a lexical property of each verb stem in the dictionary. We may reasonably assume, then, that one of the features [1 conjugation], [2 conjugation], [3 conjugation] is a lexical feature of every verb stem. These features also play a role in derivational morphology, as shown by the following sets of related forms: first conjugation tolerAr, 'tolerate' ,

tolerAncia, 'tolerance' tolerAble, 'tolerable'
tolerAnte, 'tolerant'; second conjugation creEr,
 'believe', creEncia, 'belief', creible, 'cred-
 ible', crevEnte, 'believer'.

It has been assumed all long in this study
 (...) that theme vowels do not appear as such in
 the lexicon. Rather, they are "spelled out" by
 a readjustment rule with the effect of (72):

$$(72) \quad \emptyset \rightarrow \left\{ \begin{array}{l} a / [1 \text{ conj}] \\ e / [2 \text{ conj}] \\ i / [3 \text{ conj}] \end{array} \right\} + \text{---} \#]_V$$

The results of (72) are illustrated in (73):

$$(73) \quad [v \# [1 \text{ conj}] \#]_V \quad [v \# [2 \text{ conj}] \#]_V \\ [v \# am + a \#]_V \quad [v \# com + e \#]_V$$

The formulation of (72) leaves much to be de-
 sired. For instance, it is interesting that a
 the least marked vowel, is the theme vowel of
 the least marked conjugational class, namely, the
 first (which is the largest class by far, the most
 regular, and the only productive one, with all
 new verbs invariably being assigned to it). One
 has no clear intuition about which of the two re-
 maining conjugations is the more heavily marked.
 Possibly, then, a hierarchy among the conjugations
 exists that can be stated as $[+ \text{Verb}] \rightarrow [1 \text{ conju-}$
 $\text{gation}] \rightarrow [-1 \text{ conjugation}] \rightarrow [3 \text{ conjugation}]$,
 where $[-3 \text{ conjugation}]$ corresponds to $[2 \text{ conju-}$
 $\text{gation}]$. In terms of markedness, each verb stem
 could be listed in the lexicon with one of the
 features $[u \text{ conjugation}]$, $[+3 \text{ conjugation}]$,
 $[-3 \text{ conjugation}]$, where $[u \text{ conjugation}]$ means no
 mark at all. By marking convention, stems un-
 marked for conjugational class would be assigned
 the features $[+1 \text{ conjugation}]$, $[-3 \text{ conjugation}]$,
 verbs marked $[+3 \text{ conjugation}]$ or $[-3 \text{ conjugation}]$
 would be assigned $[-1 \text{ conjugation}]$. Thus verb
 stems of the first conjugational class, that is,
 the vast bulk of verb stems in the lexicon, would
 require no mention of conjugational class features
 in their lexical entry. Then (72) could be re-
 placed by (74)

$$(74) \quad \emptyset \rightarrow \left[\begin{array}{l} + \text{voc} \\ \alpha \text{ low} \\ \alpha \text{ bach} \\ \beta \text{ high} \end{array} \right] / \left[\begin{array}{l} \alpha 1 \text{ conj} \\ \beta 3 \text{ conj} \end{array} \right] + \text{---} \#]_V$$

The content of the preceding paragraph is of
 course sheer speculation, and (74) approaches phono-

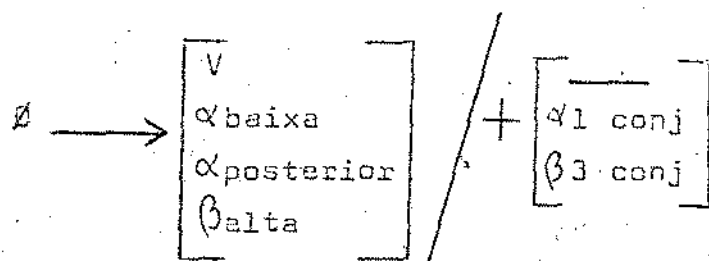
logical legerdemain. There are, however, substantive issues involved. It is clear at least that theme vowels must be supplied by rule rather than being fully specified in the lexicon. To support this assertion, one can give, in addition to vague claims involving some notion of "simplicity", the following argument. There are just three conjugational classes and hence three different theme vowels. This fact is not captured if each theme vowel is fully specified in the lexicon. That is, if theme vowels are fully specified lexically, then it is a lexical idiosyncrasy of each verb stem that its theme vowel is not /u/ or /o/. This is simply wrong. (gri-fo nosso) -- (98-100)

44. A representação fonética para o espanhol não é aqui mencionada pelos motivos expostos na nota 14 deste capítulo.

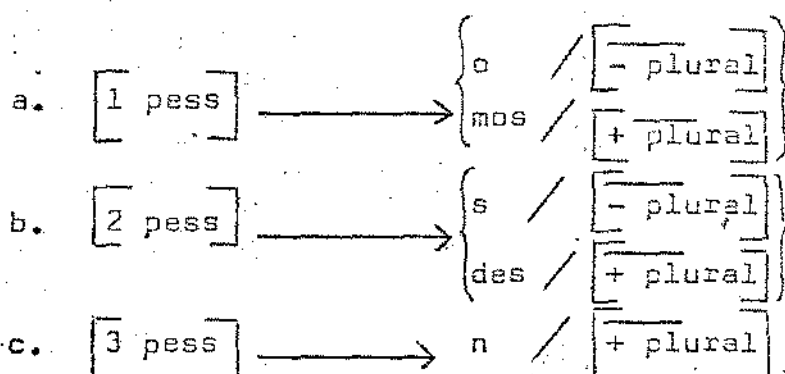
LISTA DE REGRAS

(Enumeradas por ordem de aplicação)

(1) VOGAL TEMÁTICA



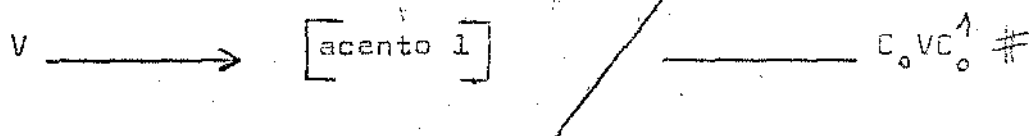
(2) FORMATIVOS GRAMATICAIS NÚMERO-PESSOAIS



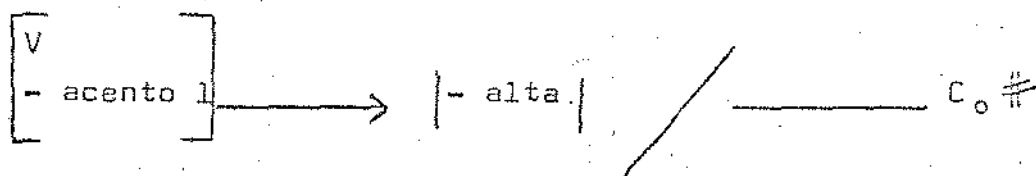
(3) TRUNCAMENTO



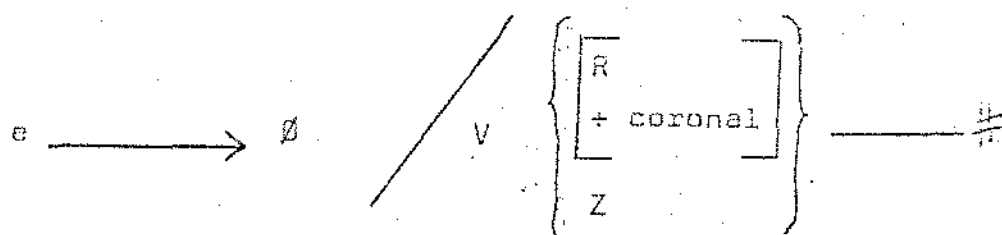
(4) ACENTO (em formas verbais)



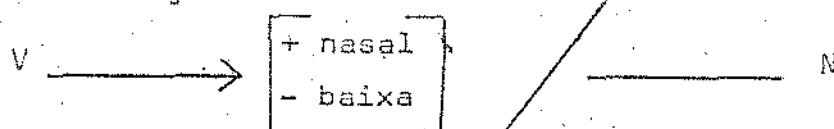
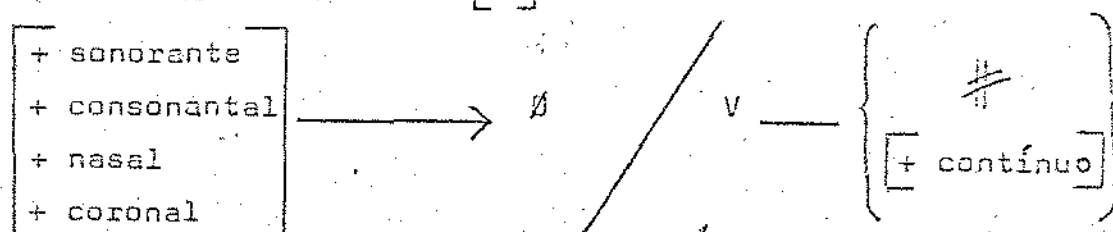
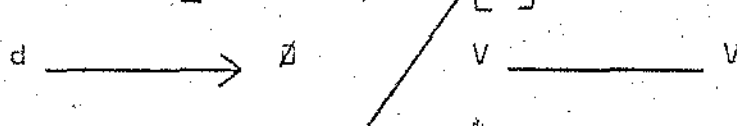
(5) ABAIXAMENTO



(6) APÓCOPE



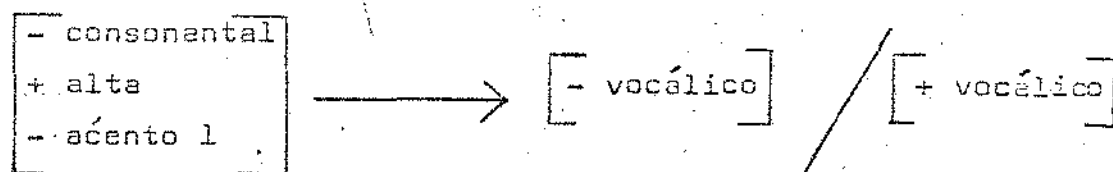
(7) NASALIZAÇÃO

(8) QUEDA DE n (formativos [+E])(9) QUEDA DE d (formativos [+E])

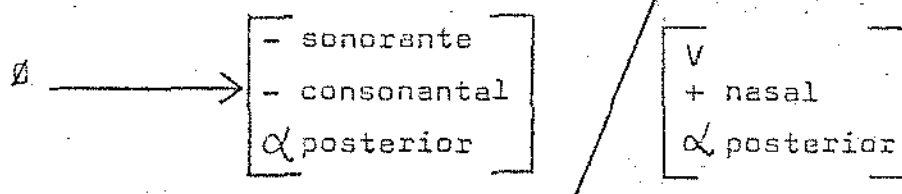
(10) LEVANTAMENTO



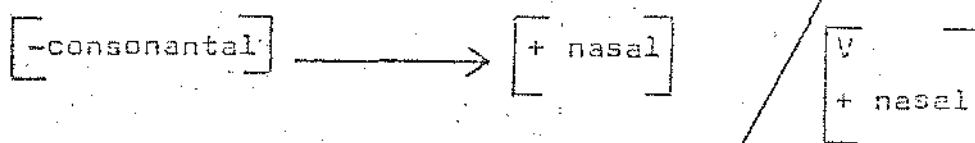
(11) FORMAÇÃO DE SEMIVOGAL



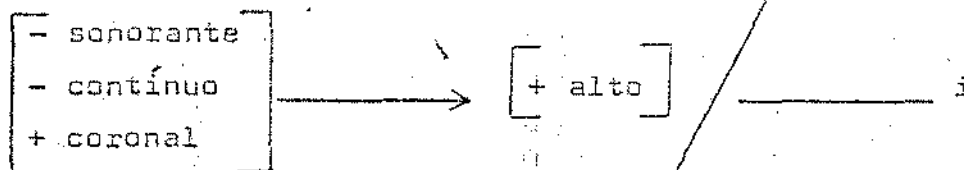
(12) INSERÇÃO DE SEMIVOGAL



(13) NASALIZAÇÃO SECUNDÁRIA



(14) PALATALIZAÇÃO



BIBLIOGRAFIA

- ANDERSON, S.R., 1969, West Scandinavian Vowel Systems and the Ordering of Phonological Rules, tese de doutoramento não publicada, MIT.
- _____. 1970, "On Grassman's Law in Sanskrit", Linguistic Inquiry, v. I nº 4.
- BACH, E., 1968, "Two Proposals Concerning the Simplicity Metric in Phonology", Glossa, 2:128-49.
- BRANE, M., 1969, "On the Abstractness of Phonology", (não publicado). MIT.
- BUHL, W.E., 1960, Time, Tense and the Verb, University of California Press, Berkeley.
- CAMARA, J.M. Jr., 1970, Estrutura da Língua Portuguesa, Editora Vozes Limitada, Petrópolis, R.J.
- _____. 1972, History and Structure of the Portuguese, University of Chicago Press, Chicago. (trad. de A.J. Naro).
- _____. 1953, Para o Estudo da Fonêmica Portuguesa, Organização Simões, Rio de Janeiro.
- _____. 1969, Problemas da Lingüística Descritiva, Editora Vozes Limitada, Petrópolis, R.J.
- CARMO, A.N. et al., 1971, Enciclopédia Orgânica da Língua Portuguesa, Brasília Editora S/A., Vitória, E.S., 5v.
- CHOMSKY, N., 1965, Aspects of the Theory of Syntax, The MIT Press, Cambridge, Massachusetts.
- _____. 1964, Current Issues in Linguistic Theory, Mouton, The Hague.
- _____. 1967, "Some General Properties of Phonological Rules", Language 43: 102-28.
- _____. 1957, Syntactic Structures, Mouton, The Hague.
- CHOMSKY, N. e M. Halle, 1965, "Some Controversial Questions in Phonological Theory", Journal of Linguistics 1: 97-138.

- _____. 1968, The Sound Pattern of English, Harper and Row, New York, N.Y.
- CHOMSKY, N., M. Halle e F. Lukoff, 1956, "On Accent and Juncture in English", in For Roman Jakobson, Mouton, The Hague, 65-80
- COUTINHO, I. de L., 1968, Gramática Histórica, Livraria Acadêmica, Rio de Janeiro.
- CROTHERS, J., (sem data), "A Note on the Abstractness Controversy", (exemplar mimeografado).
- EASTLACK, C.L., 1961, The Morphology of the Verb in Portuguese, tese de doutoramento não publicada, University of Texas.
- _____. 1964
- FODOR, J.A. e J.J. Katz (eds.), The Structure of Language: Readings in the Philosophy of Language, Prentice-Hall, Englewood Cliffs, New Jersey.
- FOLEY, J.A., 1965, Spanish Morphology, tese de doutoramento não publicada, MIT.
- HALLE, M., 1964, "On the Bases of Phonology", in: Fodor e Katz, 1964: 324-33. (trad. in: Lemle e Leite, 1978).
- _____. 1962, "Phonology in a Generative Grammar", Word 18: 54-72. (reimpresso in: Fodor e Katz, 1964: 334-52).
- _____. 1959, The Sound Pattern of Russian, Mouton, The Hague.
- HARMS, R.T., 1968, Introduction to Phonological Theory, Prentice-Hall, Englewood Cliffs, New Jersey.
- HARRIS, J.W., 1969, Spanish Phonology, The M.I.T. Press, Cambridge, Massachusetts.
- HOCKETT, C.F., 1958, A Course in Modern Linguistics, The Mac Millan Co., New York.
- HUBER, J., 1933, Altportugiesisches Elementarbuch, Carl Winters Universitätsbuchhandlung, Heidelberg.
- HYMAN, L.M., 1970, "How Concrete is Phonology?", Language 46: 58 - 76.

- JAKOBSON, R., 1967, Fonema e Fonologia, Livreria Acadêmica, Rio de Janeiro (ensaios traduzidos por J.M. Câmara Jr.).
- JAKOBSON, R. e M. Halle, 1963, Fundamentals of Language, Mouton, The Hague.
- JAKOBSON, R., G. Fant e M. Halle, 1963, Preliminaries to Speech Analysis, The M.I.T. Press, Cambridge, Massachusetts.
- KATZ, J. e P. Postal, 1964, An Integrated Theory of Linguistic Descriptions, The M.I.T. Press, Cambridge, Massachusetts.
- KIPARSKY, P., 1968, (exemplar mimeografado), "How Abstract is Phonology?", Indiana University Linguistics Club.
- KISSEBERTH, C.W., 1969, "On the Abstractness of Phonology : the evidence from Yawelmani", Papers in Linguistics 1-2, Florida State University, Tallahassee.
- KLIMA, E.S., 1959, "Negation in English" in: Fodor e Katz 1964: 246 - 323.
- KOUTSCUDAS, A., G. Sanders e C. Noll, 1971, (exemplar mimeografado), "On the Application of Phonological Rules", Indiana University Linguistics Club".
- LANGACKER, R.W. 1972, Fundamentals of Linguistic Analysis, Harcourt Brace Jovanovich, Inc., New York.
- _____. 1967-8, Language and its Structure, Harcourt, Brace & World Inc., New York.
- LEMLE, M., 1963, "Nota sobre os Alofones Surdos das Vogais na fala do Rio de Janeiro", comunicação apresentada à VII. Reunião Brasileira de Antropologia, São Paulo. (Nova versão publicada em Estudos Lingüísticos, Revista Brasileira de Lingüística Teórica e Aplicada, São Paulo, 1966, vol I nº 2, 33-4).
- LEME, M. e Y. Leite (eds.), 1970, Novas Perspectivas Lingüísticas, Editora Vozes Limitada, Petrópolis, R.J.

LENZ, R., 1944, La Oración y Sus Partes, Nascimento, Santiago, Chile.

LEED, C., 1970, Problems in Catalan Phonology, (Studies in Linguistics and Language Learning, vol. VIII), University of Washington Press, Seattle, Washington.

MALMBERG, B., (ed.), 1970, Manual of Phonetics, North - Holland Publishing Company, Amsterdam.

NARD, A.J., 1971 "The History of e and o in Portuguese", Language 47 nº 3.

PONTES, E., 1965, Estrutura do Verbo no Português Coloquial, tese de Mestrado, Gráfica da Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais, Belo Horizonte.

POSTAL, P.M., 1968, Aspects of Phonological Theory, Harper and Row, New York.

RALLIDES, C., 1971, The Tense Aspect System of the Spanish Verb, Mouton, The Hague.

SACIUK, B., 1970, "Some Basic Rules of Portuguese Phonology", in: Studies Presented to Robert E. Lees by his Students, Sadock, J.M. e A. L. Vanek (eds.), Linguistic Research Inc., Edmonton - Champaign.

SCHANE, S.A., 1968, French Phonology and Morphology, The M.I.T. Press, Cambridge, Massachusetts.

STOCKWELL, R.P., J.D. Bowen e J.W. Martin, 1965, The Grammatical Structures of English and Spanish, University of Chicago Press, Chicago.

TROUBETZKOY, N.S., 1970, Principes de Phonologie, Éditions Klincksieck, Paris.