

Transdutor de estados finitos para o reconhecimento da nasalidade na pronúncia da variedade potiguar

Cid Ivan da Costa Carvalho¹

¹Campus Caraúbas – Universidade Federal Rural do Semiárido (UFERSA)
RN 233, CEP: 59700-000 - Caraúbas-RN – Brasil

cidivanc@gmail.com

Abstract. *This paper presents a finite state transducer for transcription of the nasality to pronunciation of potiguar variety. Therefore, literature searches were made and implementation through the FOMA library. After running on a random sample of the Corpus CENTENFolha, the system showed a high performance index, with 97% hit for nasal vowels and 94% accuracy for nasality.*

Resumo. *Este trabalho apresenta um transdutor de estados finitos para a transcrição da nasalidade na pronúncia da variedade linguística potiguar. Para isso, foram feitas pesquisas bibliográfica e implementação por meio da biblioteca no foma. Após a execução em uma amostra aleatória do Corpus CENTENFolha, o sistema apresentou alto índice de desempenho, sendo que 97% de acerto para vogais nasais e 94% de acerto para a nasalidade.*

1. Introdução

Os transdutores de estados finitos são dispositivos capazes de relacionar uma cadeia de entrada a uma cadeia de saída. Apresentamos aqui um módulo do transdutor Potigrafone – Conversor de grafema para a pronúncia da variedade linguística potiguar - que relaciona os símbolos gráficos do português com os símbolos fonéticos do *Speech Assessment Methods Phonetic Alphabet* – SAMPA, para a transcrição da nasalização - as vogais nasais e do fenômeno da nasalidade. Esse sistema foi desenvolvido na linguagem *foma*, biblioteca *open source* em C para o processamento de linguagem natural, veja Hulden (2009).

Os sistemas que fazem essa relação são rotulados pela sigla *G₂P* (*grapheme to Phoneme*) podem ser construídos como um transdutor, pois eles executam a transcrição de grafema para fonema, ou seja, convertem uma sequência de grafema em uma sequência de símbolos fonológicos e/ou fonéticos que representam determinada língua ou variedade linguística.

Há alguns desses conversores para o português, mas apenas mencionaremos três: o sistema híbrido Grafone, desenvolvido por Veigas, Candeias e Perdigão (2011) para português europeu (PE). Este último está disponível no site <http://www.co.it.pt/~labfala/g2p/>; e os sistemas: *Petrus* o qual faz a transcrição do grafema para a variedade paulista - disponível em: <http://www.nilc.icmc.usp.br/petrus> - desenvolvido por Marquiefável, Bokan e Zavaglia (2014), e o *Nhenhem1.0*, desenvolvido por Vasilévski (2008), para o português brasileiro (PB).

Aqui apresentamos um dos principais módulos que integra o sistema que executa a transcrição de grafema para a variedade potiguar, nasalização. Este módulo executa a

transcrição das vogais nasais, como na palavra "campo", e faz a inserção do traço de nasalidade nas palavras em que o contexto escrito possui aspectos de vogal oral, mas é pronunciado na variedade potiguar como uma vogal nasal, por exemplo, a palavra "ama".

Outro aspecto que devemos considerar é que os símbolos fonéticos utilizados na transcrição das palavras são símbolos do SAMPA, que é um sistema de escrita fonética legível por computadores, usa um conjunto de caracteres do código ASCII (*Código Padrão Americano para o Intercâmbio de Informação*) de 7 bits e foi desenvolvido a partir de um mapeamento (codificação) dos símbolos do *International Phonetic Alphabet* - IPA.

Este trabalho está estruturado em quatro seções: a primeira apresenta a relação que há entre as línguas formais, as expressões regulares e os transdutores de estados finitos; a segunda expõe sobre a nasalização na língua portuguesa distinguindo as vogais nasais do fenômeno da nasalidade ocorrida em alguns contextos de vogais orais; a terceira mostra a construção do módulo por meio da biblioteca *foma* e a quarta apresenta o desempenho desse módulo na transcrição da nasalização para pronúncia da variedade potiguar.

2. Língua formal, expressão regular e transdutor

O termo língua é usado aqui no sentido geral para se referir a um conjunto de cadeias, também chamado de *string*, ou seja, "um conjunto formado por sequências resultantes da concatenação de elementos extraídos de um conjunto de símbolos, chamado alfabeto ou sigma". (ALENCAR, 2011, p.19). Para melhor compreensão desse conceito, suponha os elementos do conjunto $\Sigma = \{b, n, c, a, o\}$, que podem ser repetidos e concatenados entre si e com outros elementos do conjunto; então, a partir de Σ pode ser construída uma língua formal L_1 , a qual pode ter como palavras *baanco!*, *banco!*, etc.

Esse autor acrescenta que uma língua formal é caracterizada pela "enumeração exaustiva de seus elementos" e pela "especificação de 'um critério de pertinência que é satisfeito por todos os elementos do conjunto e somente por esses elementos'". (IDEM, p.21). A língua L_1 pode ser definida por meio das *expressões regulares* $R_1 = \{b, a, a, a, n, c, o\}$ ou $R_2 = b a^+ n c o$. Em R_1 , a língua é definida pela enumeração dos elementos e, em R_2 , por critérios de pertinência.

A língua L_1 e as expressões regulares R_1 ou R_2 podem ser compiladas por meio de uma *rede de estados finitos*, ou **transdutor**. A figura 1, abaixo, também chamado de *grafos de transição*, mostra a rede de *estados finitos* produzida pelas expressões regulares.

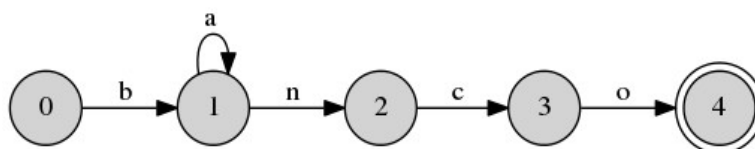


Figura 1. Grafo de transição de L_1

Esse transdutor mostra que a cadeia de entrada é idêntica a cadeia de saída. O conjunto de símbolos Σ forma um relação de pares ordenados de cadeias com os próprios elementos entre si. Para essa, Karttunen (2009) diz que o primeiro membro de

um par da relação é chamado de cadeia superior (*upper string*) e o segundo é chamado de cadeia inferior (*lower string*). Cada caminho de um transdutor representa um par de *string* numa relação. Para o transdutor de grafema para pronúncia, a relação da *upper string* com a *down string* são as representações gráficas são cadeias do nível *lower* e as representações fonéticas são cadeias do nível *upper*.

Beesley e Karttunen (2002) afirmam que um dos resultados fundamentais da teoria da língua formal é a demonstração de que os estados finitos de uma língua são precisamente um conjunto de línguas que podem ser descritas por uma expressão regular. Como podemos ver na figura abaixo. A língua é denotada por uma expressão regular, que é uma composição de símbolos, caracteres com funções especiais, que, agrupados entre si e com caracteres literais, formam uma sequência, uma expressão; e codificada em uma rede de estados finitos.

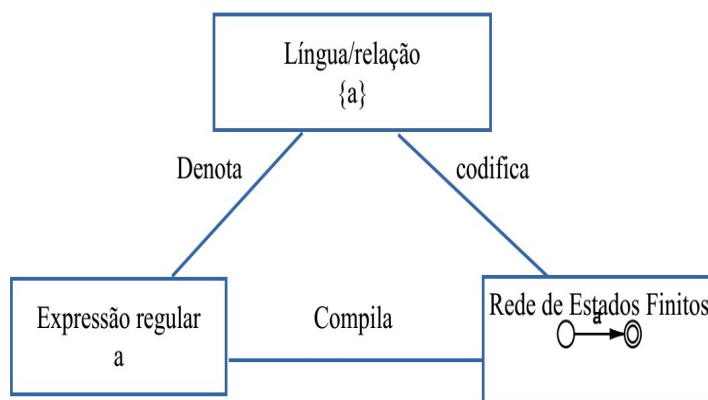


Figura 1. Relação entre a língua forma, expressão regular e transdutores

3. As vogais nasais e o fenômeno da nasalidade

O fenômeno da nasalidade, no português falado no Brasil, é um dos fenômenos sensível aos limites da estrutura silábica. “A nasalidade de uma vogal ocorre quando uma vogal tipicamente oral é seguida por uma das consoantes nasais: [m, n, nh].” (SILVA, 2014, p. 93). Esse fenômeno ocorre com maior frequência nas variedades linguísticas do Nordeste e é mais perceptível auditivamente com a vogal central baixa [a], todavia ocorre com as outras vogais.

Esse fenômeno ocorre, por exemplo na palavra "tomate" [to~mati], onde uma vogal oral transforma-se em vogal nasalizada quando está diante de uma consoante nasal na sílaba seguinte, ou seja, a vogal assimila a nasalidade por conta da abertura antecipada do véu palatino. Esse fenômeno está formalizado na seguinte regra fonológica:

$$V \rightarrow [+nasalizada] / _ .(C) \\ +Nas$$

Figura 3. Regra de inserção de traço de nasalidade

As vogais nasais são representadas, na escrita, pelos dígrafos vocálicos *am*, *an*, *em*, *en*, *im*, *in*, *om*, *on*, *um* e *um* e poucas são representadas por grafemas com diacríticos, "ã" e "õ". Nesses dígrafos, as letras "n" e "m" em posição pós-vocálica representam a nasalização das vogais, no entanto, em início de sílaba, elas são consideradas consoantes nasais.

Para fazer estimativas precisas sobre a construção silábica da nasalização das vogais, devemos considerar que o fenômeno da nasalidade, tanto no comportamento fonológico quanto na aplicação empírica, requer sutil diferenciação dos padrões silábicos no que diz respeito à relação entre *input* e *output* do sistema. Essa diferenciação é muito importante, pois pode ocorrer que um mesmo *input* tenha mais de um *output*. O mesmo padrão silábico pode gerar *outputs* incorretos. Além disso, as consoantes nasais podem sofrer alterações quando ocorrem entre vogais. Nesse situação, o sistema pode eleger como saída, para as letras "n" ou "m", os sons consonantais ou o traço de nasalização da vogal.

No próximo tópico, será mostrada a implementação do módulo na biblioteca foma que faz a transcrição da nasalização para a pronúncia da variedade linguística potiguar, onde os grafemas "n" e "m" pós-vocálicos são relacionados ao traço de nasalização e, este mesmo traço é inserido entre uma vogal precedida pelas consoantes [n], [m] e [J], distinguindo, assim, a posição silábica de ataque e coda para esses símbolos gráficos.

4. Implementação por meio do Foma

Na a implementação do sistema, utilizamos o símbolo til "˜" do alfabeto fonético SAMPA para representar a nasalização vocálica e, também, partimos do pressuposto de que: (1) C é o conjunto das consoantes; (2) V é o conjunto das vogais e (3) L é a concatenação de CV dos símbolos gráficos do português. Desse modo, definimos a seguinte expressão regular por meio do foma:

$$\text{regex } [C^* V^+ C^*] @-> \dots \text{ "˜" } || _ C V;$$

A tabela abaixo mostra os principais conceitos dos operadores utilizados no foma, veja Hulden (2009).

Tabela 1. Os principais operadores para o uso nas expressões regulares do Foma

A*	Zero ou mais vezes
A+	Uma ou mais vezes
A B	O espaço em branco representa a concatenação
A @-> B L_R	A mais longa substituição da esquerda para a direita.
...	Concatenação dos elementos
;	Fim da expressão regular
A.o.B	Composição
down	Verifica a cadeia de saída
up	Verifica a cadeia de entrada

Essa expressão regular gera todas as palavras da língua portuguesa que tenham as estruturas silábicas: V, CV, VC, CVC e separa as sílabas com o ponto “.”. Podemos verificar a geração das palavras por meio do comando *down* do *foma*. Vemos a constituição da relação do conjunto de formas *subjacentes*; as palavras de entrada; o uso da regra contextual que manipula as formas subjacentes para produzir as formas de *superfície* permitida pela língua. Isso pode ser observado mais claramente na tabela 2, onde o padrão silábico das entradas está na segunda coluna, a regra do separador silábico na terceira coluna e a língua resultante da concatenação de CV na quarta coluna.

Tabela 2. Relação de entrada e saída do sistema

Comando	Entrada	regra	Saída	palavra
down	CVCV	Inserir “.” antes da \$	CV.CV	ca.ma
	CVCCV		CVC.CV	cam.po
	VCCVC		VC.CVC	an.tes

Como foi visto no tópico anterior, para diferenciarmos as vogais nasais da nasalidade, é preciso que haja um contexto silábico que sirva como elemento norteador da cadeia linguística e contribua para a geração de transcrições bem formadas para a variedade. A construção desse proto silabificador gera as palavras separando as sílabas por meio de um ponto. Partindo dessa relação, as regras fonológicas são aplicadas às formas superficiais transformando-as e gerando novas formas de representações até o término do processo derivacional, quando se tem a forma transcrita, ou seja, aplica-se uma regra em cada processo.

As letras nasais “n” e “m” podem ser apresentadas por um ou mais símbolos fonéticos, dependendo do contexto, ou seja, a representação subjacente depende do contexto superficial. Por exemplo, a letra “n” pode representar a consoante nasal [n], quando vem em posição de ataque no início ou no meio da palavra, no entanto, a nasalização de uma vogal só ocorre quando ela vem em posição pós-vocálica. Assim, a implementação da regra a seguir distingue uma vogal nasal de uma consoante nasal.

$$(1) \quad \text{Nas} \rightarrow \text{"~"} \parallel \text{V_} \text{"."};$$

Essa regra mostra que a relação das consoantes nasais possui o traço de nasalização <Nas, ~>, antes do separador silábico. As consoantes nasais pós-vocálicas têm como representação subjacente o símbolo de nasalização. Trazendo os exemplos mencionados acima, podemos dizer que todas as letras nasais em posição pós-vocálicas são representadas pelo til “~” e permanecem inalteradas quando estiverem em posição de ataque silábico, seja no meio ou no início da palavra. Como podemos ver na tabela 3.

Tabela 3. relação da forma subjacente com a forma subjacente

Subjacente	ka.ma	ka~.pu	a~.tis
Regras	Consoantes nasais serão traços de nasalização		
Superfície	cama	campo	antes

O fenômeno da nasalidade ocorre com as vogais orais que assimilam o traço nasal da consoante nasal da sílaba subsequente, como foi ilustrado na figura 3. Esse fenômeno pode ser implementado pela expressão regular a seguir:

(2) $[..] \rightarrow " \sim " || V_Nas;$

O "a" que antecede a consoante nasal na palavra "cama" da tabela 3 é pronunciada na variedade potiguar como uma vogal nasal. O uso dessas regras tem o propósito de explicitar esse fenômeno fonético. A transcrição dessa palavra recebe a inserção do traço nasal, [ka~.ma], pois satisfaz às exigências contextuais de escrita para esse fenômeno.

5. Avaliação do sistema

Esse módulo do sistema foi avaliado numa listagem de 750 palavras do *corpus* CETENFolha. Esse *corpus* é acessado por meio do sistema de busca da linguateca AC/DC (Acesso a corpos/Disponibilização de corpos), utilizando a classe JOCF. Na pesquisa, esse sistema de busca apresentou uma lista em ordem decrescente das palavras mais frequentes. O processo de criação dessa listagem consistiu em tomar as cadeias de caracteres anotadas como palavras, obedecendo aos seguintes critérios: começar com um grafema da língua portuguesa; não conter dígitos; não apresentar grafemas em maiúsculas; não conter caracteres como ".", ",", "-", etc.; e a exclusão dos nomes próprios.

Após esse processo, fizemos a contagem das palavras que possuíam as letras "n" ou "m" em posição pós-vocálica, formando um dígrafo vocálico, e em posição medial, entre vogais. Do total de palavras, foram encontrados 242 casos de consoantes nasais, dentre os quais 175 formavam vogais nasais, correspondendo a 23% - e 67 eram consoantes em posição medial, que corresponde a 6% da amostra, como mostra o gráfico na figura a seguir. Ressaltamos que as ocorrências dessas letras no início de palavras foram desconsideradas.

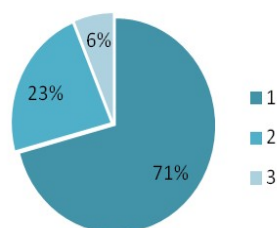


Figura 5. Regra de inserção de traço de nasalidade

Algumas palavras, no entanto, podem apresentar um ou mais fenômeno de nasalidade e/ou uma ou mais vogais nasais. Contamos, então, cada fenômeno separado na transcrição automática das palavras, para observarmos o nível de acurácia do sistema relativos às vogais nasais e à nasalidade. A figura 6 mostra a taxa porcentual do desempenho do módulo para cada um desses fenômenos. Assim, os gráficos tenta responder qual o percentual das palavras de entrada foram transcritas corretamente para cada fenômeno? O primeiro gráfico mostra que apenas 3% foram transcritas de forma errada. Esses erros ocorreram em palavras que apresentam o encontro das duas consoantes nasais, como em "amnésia" [a~nEzja]. O segundo gráfico mostra que apenas 6% receberam a inserção desse traço erroneamente. Os erros mais comuns para esse

fenômeno ocorreram nas palavras terminadas em “mente”, como em "respectivamente" [respectiva~menti], onde não ocorre nasalidade nesse contexto para essa variedade.

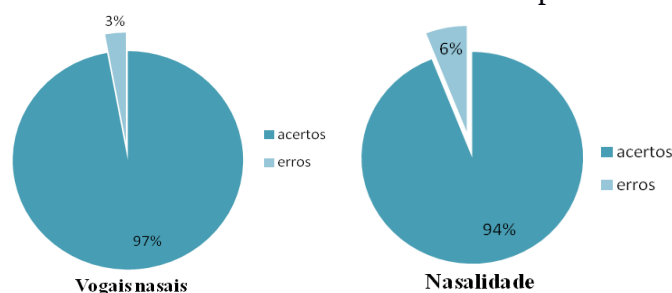


Figura 6. Ocorrência das vogais nasais e da nasalidade

6. Considerações gerais

Apresentamos aqui o módulo do transdutor potigrafone que transcreve os grafemas do português com os símbolos do alfabeto fonético SAMPA para fala potiguar das vogais nasais e do fenômeno do nasalidade. Ele foi desenvolvido na linguagem *foma*, biblioteca *open source* em C, para o processamento de linguagem natural.

Esse módulo do sistema apresentou alto índice de desempenho para esses fenômenos, como mostra figura 6, mas há alguns erros muito pontuais devem ser corrigidos para o aperfeiçoamento do sistema. Consideramos que esse módulo poderá ser integrado aos sistemas de reconhecimento ou de síntese de voz para essa variedade linguística, mas que, no momento, o nível de pesquisa se volta apenas para a transcrição fonética.

7. Referências

- Alencar, L. F. de. (2011) “Línguas formais, gramáticas e autômatos no processo automático das palavras”, In: Alencar, L. F e Othero, G. de A. *Abordagens computacionais da teoria da gramática*. Campinas-SP: Mercado de Letras, p. 13-76.
- SILVA, T. C. (2014). *Fonética e fonologia do português: roteiro de estudos e guia de exercício*. 10. ed. São Paulo: Contexto.
- Beesley, K. R. e Karttunen, L. (2002). *Finite-State Morphology*. Xerox Tools and Techniques. Cambridge University Press, Cambridge, England.
- Hulden, M. (2009). “Foma: a finite-state compiler and library.” *Proceedings of the EACL*, Athens, Greece, 3 April, p. 29–32.
- Marquiasável, V. e Zavaglia, C. (2011) "Transcrição fonética automática para lemas em verbetes de dicionários do Português do Brasil" in: *Proceedings of the 8th Brazilian Symposium in Information and Human Language Technology*, Cuiabá - MT, Brasil, outubro 2011, p. 154-158. disponível em: http://nilc.icmc.sc.usp.br/til/stil2011_English/stil/artigos/Short/STIL2011_SP1.pdf. Acesso em: 10 de agosto de 2015.
- Vasilévski, V. (2008) "Construção de um sistema computacional para suporte à pesquisa em fonologia do português do Brasil". *Tese de doutorado* - Pós-graduação em Linguística da Universidade Federal de Santa Catarina.
- Veiga, A., Candeias, S., Perdigão, F. (2011) "Conversão de Grafemas para Fonemas em

Português Europeu – Abordagem Híbrida com Modelos Probabilísticos e Regras Fonológicas." In: *Linguamática*, Dezembro 2011.

Karttunen, L. (2009) "Finite-State technology" In: MITKOV, Ruslan. *The oxford handbook of computational linguistics*. New York: Oxford University Press.