Software para o Estudo da Língua Grega Clássica: formalismo subjacente a mecanismos adaptativos

E. G. Cavalcanti, USP, Brasil, Z. M. Zapparoli, USP, Brasil

***Resumo* — Esta pesquisa tem por objeto de estudo a flexão nominal do dialeto ático em ambiente digital. Os objetivos principais são: 1) construção de programa de computador que viabilize a flexão nominal, fornecendo a análise da palavra digitada; 2) criação de corpus do dialeto ático; 3) criação de mecanismo de busca direta no corpus do TLG, Thesaurus Linguae Graecae. São três as contribuições: a primeira – estritamente lingüística –, pela apresentação das desinências nominais sinteticamente; a segunda – interdisciplinar, na convergência das novas tecnologias digitais com as letras clássicas –, pela disponibilidade da análise morfológica em ambiente digital de forma integral e simultânea; a terceira – no emprego das Tecnologias Adaptativas –, pela geração de uma Base de Conhecimentos com os recursos complementares capazes de fornecer a combinação correta do grego ático.**

***Palavras-chave:* flexão nominal; língua grega; dialeto ático; novas tecnologias digitais; programa; lingüística; cálculo morfológico; dispositivos adaptativos; reconhecimento automático de padrões.**

I. Introdução

ESTE trabalho tem por objetivo mostrar os resultados da pesquisa de doutorado, em relação à constituição dos léxicos e à construção da flexão nominal, ambos do dialeto ático, através de *softwares* especificamente criados para esses fins.

Em primeiro lugar, apresentamos o objeto de pesquisa, as justificativas e os objetivos do projeto. Em segundo lugar, discorremos a respeito do desenvolvimento digital, esclarecendo os aspectos relativos ao ambiente de programação e ao banco de dados utilizados. Mostramos, em terceiro lugar, como se realizou a constituição do *corpus* de estudo, em seus diversos desdobramentos. Em quarto lugar, expomos o mecanismo da flexão nominal em ambiente digital: cálculo das combinações – programa *Nomes*. Em último lugar, falamos sobre as tecnologias adaptativas no sistema nominal do grego ático, especificamente no que diz respeito ao reconhecimento automático de padrões do sistema nominal do grego antigo.

II. Objeto de estudo, justificativa e objetivos

Este trabalho tem por objeto de estudo a flexão nominal do dialeto ático – uma das variantes dialetais do grego antigo, juntamente com o jônico, o dórico e o eólico. O interesse em processar o dialeto ático em ambiente digital surgiu em 2001, nas aulas de grego clássico do professor Dr. Henrique Graciano Murachco – Departamento de Letras Clássicas –, nos cursos de extensão da Universidade de São Paulo.

O fato de o método do professor Henrique não privilegiar o modelo das gramáticas tradicionais, considerado por ele impositivo e prescritivo, mas *pensar* as formas lingüísticas – nominais e verbais –, levando em consideração as suas partes fundamentais – a saber, o tema (parte fixa, sede do significado virtual, absoluto ou abstrato) e as desinências (que estabelecem a função, a pessoa, o número, a voz, o modo e o aspecto) – e os processos fonéticos que integram a sua *história*, criou-nos a convicção de sua coerência metodológica, visto desaparecerem as irregularidades morfológicas e, portanto, a idéia de que era possível flexionar as formas em um programa de computador. Bastaria, para isso, criar um banco de dados com todas as desinências nominais e verbais, extrair a desinência da palavra a ser pesquisada e submetê-la a um processo comparativo com as desinências do banco e, desse modo, gerar as combinações possíveis. Criamos, então, um protótipo digital da flexão ática e, para este trabalho, decidimos desenvolver o módulo relativo à flexão nominal.

Embora concordando com que as gramáticas gregas são impositivas, não podemos negar o papel de importância indiscutível que elas exerceram no ensino da língua grega. Recorremos, no nosso estudo, inúmeras vezes a elas, ora coletando informações, ora confrontando-as em pontos duvidosos ou lacunosos, mas *sempre as consultando*. São nelas que, bem ou mal, encontramos o *depósito das formas.* Ainda que façamos uso de programas de computador em nossas pesquisas, nossa intenção não é produzir um discurso de refutação, de desconsideração ou negação da interferência do outro, mas de complementação. Acreditamos que o saber é coletivo, interdiscursivo, dialógico, feito por várias mãos e povoado por várias vozes. Reconhecemos as limitações, mas, também, o valor histórico das gramáticas.

O avanço das novas tecnologias digitais em todos os ramos da vida humana coloca, entretanto, desafios intransferíveis às pesquisas lingüísticas. Não obstante as gramáticas tradicionais tenham desempenhado papel relevante na história do ensino, é imprescindível, hoje, que o conhecimento lingüístico esteja integrado ao conhecimento digital.

Há um *déficit lingüístico* muito grande em estudos de língua grega no Brasil que utilizam os recursos computacionais de forma integrada. Não há ferramenta informática – grego-português – que forneça *simultaneamente* diversas informações lingüísticas sobre as formas nominais. Por causa disso, entendemos que a justificativa inicial de nosso trabalho esteja relacionada ao preenchimento da lacuna nos estudos das línguas clássicas. A nossa contribuição relaciona-se, portanto, à criação de um programa de computador que processe informações sobre as formas nominais, isto é, realize a flexão nominal, estabelecendo, *simultaneamente,* o tema, o caso, o gênero, o número, *o conjunto das combinações* derivadas da *análise e do cálculo morfológicos*, quadros paradigmáticos de flexão, entrada no dicionário e busca rápida em *corpus extenso* com margem de erro desprezível, o qual possibilite o confronto com as formas criadas pelo computador, de forma a apontar para a validade ou não das combinações. Isso realizado, diríamos que à ferramenta caberia o processamento de todas as informações e ao estudioso, a tarefa de compreender os conceitos e interpretar os resultados.

Outro motivo relaciona-se à eficácia fornecida pelas tecnologias digitais. Todos nós sabemos como é penoso "decifrar" o nominativo singular dos nomes da terceira declinação (nomes com temas em consoante, soante, semivogal) ou das formas contratas, no grego antigo. Por causa dos acidentes fonéticos – que provocam as aparentes irregularidades morfológicas –, a busca da entrada nos dicionários torna-se, muitas vezes, desanimadora. É necessário que se tenha pleno domínio dos processos fonéticos, para efetuar a decomposição das formas e, assim, encontrar o nominativo singular e, portanto, a entrada no dicionário. Isso implica em investimento de *muito* tempo em exercícios de tradução para se adquirir o *know-how* satisfatório, o que provoca a desistência, não raras vezes, do estudante da língua grega. Por outro lado, se armazenássemos em um banco de dados todas as desinências nominais, com seus respectivos processos fonéticos, e transformássemos a "rotina" humana de decodificação nominal em um programa de computador, necessariamente o tempo da análise diminuiria absurdamente.

Postulamos, portanto, inequivocamente, o uso das novas tecnologias digitais nos estudos lingüísticos, não como algo *estranho* que se agrega, mas como elemento constitutivo do processo de produção do conhecimento lingüístico. Essa postulação está filiada, como vimos, a uma tendência universal irreversível de informatização generalizada das rotinas humanas que exigem repetição, precisão, volume e rapidez no processamento; tendência, aliás, presente nos processos científicos e acadêmicos de produção do conhecimento dos grandes centros de excelência.

Em coerência com o exposto acima, objetivamos, com este trabalho:

1) construção de programa de computador que viabilize a flexão nominal, fornecendo a análise da palavra digitada, isto é: o tema nominal – em vogal, consoante, soante, semivogal; o caso da forma digitada pelo usuário – nominativo, vocativo, acusativo, genitivo, dativo, locativo e instrumental; o gênero – masculino, feminino, neutro; o número – singular e plural; o particípio – *infectum, aoristo* e *perfectum,* em suas vozes ativa, média e passiva; tabela com todas as combinações propostas pelo programa, com seus respectivos quadros de flexão; entrada no dicionário – nominativo singular; quadros paradigmáticos para eventual confronto com as combinações propostas pelo programa;

2) criação de *corpus* do dialeto ático que permita a geração de um léxico em banco de dados, com indexação diferenciada dos itens lexicais, indicando o autor, a obra e a linha, visando verificar a existência das formas criadas pelo programa – busca indireta – e fornecer subsídios para eventual análise lingüística, através do cálculo dos valores lexicais em cada autor e em cada item lexical. Para isso, usaremos a noção *desvio reduzido* da Estatística Paramétrica, explicitada no método matemático-estatístico-computacional de análise de textos de André Camlong, para a confecção das T.V.Ls. (Tabelas de Valores Lexicais);

3) criação de mecanismo de *busca direta* no *corpus* do TLG, *Thesaurus Linguae Graecae* (Centro de Pesquisa da Universidade da Califórnia, *Irvine*, fundado em 1972, que digitalizou textos literários escritos em grego, desde Homero – século VIII a.C. – até a queda do Império Bizantino em 1453, totalizando 3800 autores, 12000 obras, 99 milhões de palavras). A busca direta nos textos tem por finalidade única verificar, de forma rápida e ampla, a existência das formas propostas pelo programa. A diferença entre as duas buscas – a indireta e a direta – reside no tamanho do universo lingüístico pesquisado e na forma da organização dos dados: a indireta realiza a pesquisa *no léxico* de *11 autores*, dentro de um *banco de dados,* e a direta busca a forma *nos textos* dos *3800 autores*. A intenção de alargar o universo de busca é pôr em prova, através de um *corpus* "gigante", as combinações das formas propostas. Se houver convergência entre as combinações propostas e a busca realizada no *corpus*, isto é, se a busca estornar resultado positivo, isso indica que as formas propostas pelo programa procedem;

4) reorganização das desinências nominais em *conjuntos de desinências*, diferentemente da organização apresentada pelas gramáticas;

5) criação das tabelas de flexão dos particípios. As gramáticas apresentam apenas o nominativo masculino, feminino, neutro singular, em cada aspecto verbal, deixando de desenvolver os demais casos;

6) apresentação do processo fonético que deu origem à determinada forma nominal.

III. DESENVOLVIMENTO: AMBIENTE E BANCO DE DADOS

A escolha do ambiente de programação implicou algumas dificuldades. A primeira delas diz respeito à grande variedade de linguagens de programação existentes no mercado. Precisávamos de uma linguagem que fosse produtiva, confiável e que tivesse acesso a vários bancos de dados. A segunda, à não-existência de códigos abertos (*open source*) ou projetos disponíveis semelhantes para que pudéssemos ter um ponto de partida na construção de um projeto de flexão nominal do dialeto ático. A terceira, ao suporte técnico à linguagem escolhida

Dentre as linguagens do mercado, optamos por trabalhar com as que eram orientadas a objetos. Dentre elas, destacam-se o C++ e o Delphi, em sua versão 7, que têm, como linguagem nativa, o Pascal orientado a objetos. Decidimos pelo Delphi, pelo fato de ser um ambiente de fácil manipulação e de alta produtividade, além de ter suporte técnico disponível, fornecido por várias empresas, no Brasil (filiamo-nos ao Clube do Delphi, sediado na cidade de Avaré, interior do estado de São Paulo).

A escolha do banco de dados também não foi simples. Deveria ser robusto, confiável, com excelente performance e gratuito. Os que se encaixavam nessas características eram: SQL Server 2005, PostGresSQL, FireBird e MySQL. Para efetuarmos a escolha, realizamos um teste de performance com a tabela do léxico do dialeto ático, com 200689 registros. A busca realizou-se sobre o último registro da tabela. O resultado foi o seguinte:

TABELA1: PERFORMANCE DOS BANCOS DE DADOS

200689 registros **Busca do último registro**					
Banco de Dados (Gratis)	***Tempo em ms***	***Performance***		***Problemas de cache***	***Nativo do Delphi 7***
SQL Server 2005	**140**	**21,34%**	**1º**	**N**	**S**
PostGresSQL	**172**	**26,22%**	**2º**	**N**	**N**
FireBird	**188**	**28,66%**	**3º**	**S**	**S**
MySQL	**656**	**100,00%**	**4º**	**S**	**N**

A escolha, a partir do teste de performance, ficou clara: SQL Server 2005. Embora sua capacidade esteja limitada a 4 GB por banco, é extremamente rápido, não dá problemas de *cache* local e pode ser acessado por componente nativo do Delphi – o componente ADO, *Active Data Objects.*

Na atual fase do projeto, utilizamos o SQL Server 2005 e o FireBird. Na etapa final, migraremos todos os dados para um único banco, isto é, o SQL Server 2005.

IV. CONSTITUIÇÃO DO *CORPUS* ÁTICO: FONTES, OBRAS, AUTORES, EXTRATO DOS TEXTOS E LÉXICO

O *corpus* que deu origem ao léxico foi formado a partir da coleta de textos na *Internet,* nos *sites* de domínio público, que disponibilizam obras de autores antigos. Basicamente, foram duas as fontes: a) *Perseus Digital Library,* em http://www.perseus.tufts.edu/; e b) *Thesaurus Linguae Graecae*, na seção aberta ao público, em http://www.tlg.uci.edu/demo.html.

Os textos extraídos da *Internet* foram processados e submetidos a uma indexação diferenciada das encontradas comumente. Para tal, criou-se uma *ferramenta digital*, de nome *Corpus,* que gerou o Léxico com: a) as freqüências de uso dos itens lexicais nas obras processadas, no todo e nas partes – autor por autor; b) o endereço de cada item lexical, contendo: nome do autor, obra e linha em que se deu a ocorrência da palavra; c) o peso (valor) lexical nas obras do *Corpus.*

Tendo em vista o objeto deste trabalho, optamos por selecionar um conjunto de autores que escreveram em grego ático. A obra completa de cada autor está em formato texto, num mesmo arquivo. Isso facilita a busca morfológica e o reconhecimento por qualquer plataforma, como *Windows* e *Linux*.

Pelo fato de utilizarmos o *corpus* do TLG nas buscas rápidas, mantemos a correspondência do nome do autor com o nome do arquivo correspondente no TLG, visando a uma uniformidade. Assim:

TABELA2: AUTORES QUE ESCREVERAM EM GREGO ÁTICO

Aeschylus trag.	**TLG0085.txt**
Aristophanes comic.	**TLG 0019.txt**
Aristoteles	**TLG 0086.txt**
Demosthenes orat.	**TLG 0014.txt**
Euripides trag.	**TLG 0006.txt**
Isocrates orat.	**TLG 0010.txt**
Lysias orat.	**TLG 0540.txt**
Plato	**TLG0059.txt**
Sophocles trag.	**TLG 0011.txt**
Thucydides hist.	**TLG 0003.txt**
Xenophon hist.	**TLG 0032.txt**

As obras dos autores que escreveram em grego ático, no nosso *corpus*, somam um total de 316.

Os textos extraídos na *Internet* passaram por um tratamento informático, através do programa *Corpus*, módulo Ferramentas, com formatação específica. Mantivemos o formato *beta code*, porém em letra minúscula. Seguem-se os extratos dos textos – obras – de ***Aeschylus*** *trag.*, e ***Aristoteles***, com o endereço lexical à esquerda, bem como um recorte do léxico:

Extrato dos textos: *Aeschylus trag.*,

Aesch.Suppl01.000001: {ΙΚΕΤΙΔΕΣ} {ΧΟΡΟΣ ΔΑΝΑΙΔΩΝ} Ζεὺς μὲν ἀφίκτωρ ἐπίδοι προφρόνως
Aesch.Suppl01.000002: στόλον ἡμέτερον ναίον ἀρθέντ'
Aesch.Suppl01.000003: ἀπὸ προστομίων λεπτοψαμάθων
Aesch.Suppl01.000004: Νείλου. Δίαν δὲ λιποῦσαι χθόνα σύγχορτον Συρίᾳ
Aesch.Suppl01.000005: φεύγομεν,
Aesch.Suppl01.000006: οὔτιν' ἐφ' αἵματι δημηλασίαν
Aesch.Suppl01.000007: ψήφῳ πόλεως γνωσθεῖσαι,
Aesch.Suppl01.000008: ἀλλ' αὐτογενεῖ φυξανορίᾳ, γάμον Αἰγύπτου παίδων ἀσεβῆ
Aesch.Suppl01.000009: τ'
Aesch.Suppl01.000010: ὀνοταζόμεναι διάνοιαν.
Aesch.Suppl01.000011: Δαναὸς δὲ πατὴρ καὶ βούλαρχος

[redacted]

οἱ περιλειφθέντες τῶν Σπαρτῶν, ὡς Αἰσχύλος φησίν, ἦσαν
Aesch.Fragm11.015264: Χθόνιος, Οὐδαῖος,
Aesch.Fragm11.015265: ασιος Αἰσχύλος ομα Λακεδαιμόνιον ἀποφαίνει τὸν Αλκμᾶνα· λέγει γὰρ ἐν τοῖς
Aesch.Fragm11.015266: ' Υακινθ ἄκουσα ταν ἀηδον παρ' Εὐρώτα ταν' Αμυκλα
Aesch.Fragm11.015267: μεναι τατ τον εὐνομω ουσαναυτα
Aesch.Fragm11.015268: ἀρεταν ταν που μέλεσι ταλλαν ταν τ ' Αταρνίδα ἐν
Aesch.Fragm11.015269: · ἐν γὰρ τούτοις γράφειν ταπ ' Αλκμᾶνος τον αλ
Aesch.Fragm11.015270: ἐπιλε

Extrato dos textos: *Aristoteles*

Arist.Analy01.000001: {ΑΝΑΛΥΤΙΚΩΝ ΠΡΟΤΕΡΩΝ Α.} Πρῶτον εἰπεῖν περὶ τί καὶ τίνος
Arist.Analy01.000002: ἐστὶν ἡ σκέψις, ὅτι περὶ ἀπόδειξιν καὶ ἐπιστήμης
Arist.Analy01.000003: ἀποδεικτικῆς· εἶτα διορίσαι τί ἐστι πρότασις καὶ τί ὅρος καὶ τί
Arist.Analy01.000004: συλλογισμός, καὶ ποῖος τέλειος καὶ ποῖος ἀτελής, μετὰ
Arist.Analy01.000005: δὲ ταῦτα τί τὸ ἐν ὅλῳ εἶναι ἢ μὴ εἶναι τόδε τῷδε, καὶ
Arist.Analy01.000006: τί λέγομεν τὸ κατὰ παντὸς
Arist.Analy01.000007: ἢ μηδενὸς κατηγορεῖσθαι.

[redacted]

τοῦ χειμῶνος γίνεσθαι τοὺς ὑετούς ... ἐπιζητήσαι δ' ἄν τις,
Arist.Fragm55.117324: πῶς εἰ τὸ θέρος ξηρόν ἐστιν, ὡς προείρηκε, παρ' οἷς ξηρότατόν ἐστι τὸ θέρος,
Arist.Fragm55.117325: παρὰ τούτοις οἱ ὑετοὶ τοῦ θέρους γίνονται. ἢ αἱ μὲν ἀναθυμιάσεις καὶ αἱ τῶν
Arist.Fragm55.117326: νεφῶν συστάσεις
Arist.Fragm55.117327: οὐκ ἐκεῖ γίνονταί τε καὶ
Arist.Fragm55.117328: γεννῶνται, ἀλλ' ὑπὸ πνευμάτων τῶν ἐτησίων πρὸς τὰ ἐκεῖ ὄρη ὄντα ὑψηλὰ ἀπωθούμεναι συνίστανται, ὡς λέγει, τῆς δὲ ταχείας
Arist.Fragm55.117329: αὐτῶν εἰς ὑετὸν μεταβολῆς ἡ προειρημένη ἀντιπερίστασις αἰτία, ὡς λέγει ἐν
Arist.Fragm55.117330: τοῖς Περὶ τῆς τοῦ Νείλου ἀναβάσεως.

Recorte do léxico:

TABELA 3 – LÉXICO

	Total
Palavras	**3038958**
Itens lexicais	**200689**

Ordem	Palavra	Total	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
1	α	4	0	0	0	0	0	0	0	2	2	0	0
2	αβρων	1	0	0	0	1	0	0	0	0	0	0	0
3	αβρωνος	3	0	0	0	3	0	0	0	0	0	0	0
4	αδην	1	0	0	0	0	0	0	0	1	0	0	0
5	αγνων	5	0	0	0	0	0	0	0	0	0	5	0
200680	ζωτικην	3	0	0	2	0	0	0	0	1	0	0	0
200681	ζωτικης	1	0	0	0	0	0	0	0	1	0	0	0
200682	ζωτικον	1	0	0	1	0	0	0	0	0	0	0	0
200683	ζωτικον	3	0	0	1	0	0	0	0	1	0	0	1
200684	ζωτικωτερα	1	0	0	1	0	0	0	0	0	0	0	0
200685	ζωτικωτερον	1	0	0	1	0	0	0	0	0	0	0	0
200686	ζωτικῳ	1	0	0	0	0	0	0	0	1	0	0	0
200687	ζωτικωτερους	1	0	0	0	0	0	0	0	0	0	0	1
200688	ζωυφιον	4	4	0	0	0	0	0	0	0	0	0	0
200689	ζχ	2	0	0	2	0	0	0	0	0	0	0	0

V. O MECANISMO DA FLEXÃO NOMINAL EM AMBIENTE DIGITAL: CÁLCULO DAS COMBINAÇÕES – PROGRAMA NOMES

O CÁLCULO MORFOLÓGICO OBEDECE ÀS SEGUINTES ETAPAS:

A) ENTRA-SE COM O VOCÁBULO GREGO

Figura 1 – Entrada do vocábulo no programa Nomes

b) inicia-se um laço (*loop*) com repetição igual ao número de casos, isto é, 10.

laço 1: nominativo singular;
laco 2: vocativo singular;
laço 3: acusativo singular;
laço 4: genitivo singular;
laço 5: locativo, dativo, instrumental singular (D.L.I.);
laço 6: nominativo plural;
laco 7: vocativo plural;
laço 8: acusativo plural;
laço 9: genitivo plural;
laço 10: locativo, dativo, instrumental plural (D.L.I.).

c) abre-se a tabela com as desinências.

d) faz-se uma análise comparativa do vocábulo digitado com as desinências da tabela, da primeira até a última, caso a caso. O ponto de referência é a desinência cadastrada no banco de dados.

Vocábulo digitado:

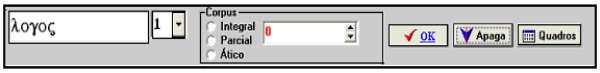

Figura 2 – Entrada do vocábulo no programa Nomes

TABELA 4 - ENTRADA DO VOCÁBULO E DESINÊNCIAS RELACIONADAS AO VOCÁBULO λογος

Singular							Plural				
ORDEM	N.	V.	A.	G.	D.L.I.		N.	V.	A.	G.	D.L.I.
1	ος	ε	ον	ου	ῳ		οι	οι	ους	ων	οις
24	ος		ον	ου	ῳ		οι		ους	ων	οις
36	ς		να	νος	νι		νες		νας	νων	σι

e) calcula-se quantos caracteres formam o vocábulo digitado. Em λογος, p. ex., temos 5 caracteres.

f) estabelece-se um novo *loop* com repetição igual ao número de caracteres do vocábulo.

g) dentro desse novo *loop*, extrai-se a desinência do vocábulo digitado, do fim para o começo, seguindo os critérios:

1. quantidade de caracteres a partir da extensão desinencial (ED), dada no banco.
2. posição do caracter, do último até o limite da ED.

P. ex.: vocábulo λογος → no banco, uma das desinências: **ος, ED=2.**

1	2	3	4	5
λ	ο	γ	ο	ς

Quantidade de caracteres: 5.

3. *loop* com repetição 5.

4. Laço 1:

■extração do caracter ς "raiz-tema" λογο-

5. Laço 2 (ED = 2):

■extração do caracter ο "raiz-tema" λογ-

6. efetua-se ou não a combinação.

Situações:

1. a desinência do banco (DB) concorda, por completo, com a desinência do vocábulo (DV), e as flexões procedem – registro 01;

2. a desinência do banco (DB) concorda, por completo, com a desinência do vocábulo (DV), e as flexões não procedem – registros 24 e 36 ;

3. a desinência do banco (DB) concorda, parcialmente, com a desinência do vocábulo (DV), não há flexões – registro 06;

4. a desinência do banco (DB) não concorda, na última posição, com a desinência do vocábulo (DV), não há flexões – registro 02.

Como resolver o empate, quando houver concordância entre DB e DV? Situações 1 e 2. Há algum indicador independente dos indicadores gramaticais? Há algum indicador independente dos indicadores gramaticais?

I. Estando no nominativo, registro 1:

1. Decomposição da desinência do vocábulo de acordo com a desinência do banco:
   a. vocábulo digitado: λογος
   b. tabela → registro 1 → desinência do nominativo: ος (extensão desinencial (ED)2)
   c. decomposição do vocábulo (ED = 2):
      i. extraindo-se o último caracter: λογος = λογο (provável raiz)
         • comparação do ς (posição final) de λογος com ς (posição final) da desinência
         • sendo iguais, continua-se a extração
      ii. extraindo-se os dois últimos caracteres: λογος = λογ (raiz)
         • comparação do ος (posição final+1) de λογος com ος (posição final + 1) da desinência
         • sendo iguais, efetua-se a soma
   d. soma do que sobrou do vocábulo digitado (raiz) com a desinência do banco:
      i. raiz, a partir da decomposição: λογ
      ii. desinência do banco: ος
      iii. soma de ambos: λογ + ος = λογος (combinação)
      iv. combinação = vocábulo digitado
2. Processam-se as combinações.

II. Estando no nominativo, registro 02:

1. Decomposição da desinência do vocábulo de acordo com a desinência do banco:
   a. vocábulo digitado: λογος

b. tabela → registro 2 → desinência do nominativo: ον (extensão desinencial (ED): 2)

c. decomposição do vocábulo (ED = 2):

i. extraindo-se o último caracter: λογος = λογο (provável raiz)

• comparação do ς (posição final) de λογος com ν (posição final) da desinência

• são diferentes. Combinação interrompida, a desinência do banco não corresponde à desinência do vocábulo.

III. Estando no nominativo, registro 06:

1. decomposição da desinência do vocábulo de acordo com a desinência do banco:

a. vocábulo digitado: λογος

b. tabela → registro 06 → desinência do nominativo: ας (extensão desinencial (ED): 2)

c. decomposição do vocábulo (ED = 2):

i. extraindo-se o último caracter: λογος = λογο (provável raiz)

• comparação do ς (posição final) de λογος com ς (posição final) da desinência

• sendo iguais, continua-se a extração

ii. extraindo-se os dois últimos caracteres: λογος = λογ (raiz)

• comparação do ος (posição final+1) de λογος com ας (posição final+1) da desinência

• são diferentes. Combinação interrompida, a desinência do banco não corresponde à desinência do vocábulo.

IV. Estando no nominativo, registro 24:

1. decomposição da desinência do vocábulo de acordo com a desinência do banco:

a. vocábulo digitado: λογος

b. tabela → registro 24 → desinência do nominativo: ος (extensão desinencial (ED): 2)

c. decomposição do vocábulo (ED = 2):

i. extraindo-se o último caracter: λογος = λογο (provável raiz)

• comparação do ς (posição final) de λογος com ς (posição final) da desinência

• sendo iguais, continua-se a extração

ii. extraindo-se os dois últimos caracteres: λογος = λογ (raiz)

• comparação do ος (posição final+1) de λογος com ος (posição final+1) da desinência

• sendo iguais, efetua-se a soma

d. soma do que sobrou do vocábulo digitado (raiz) com a desinência do banco:

i. raiz, a partir da decomposição: λογ

ii. desinência do banco: ος

iii. soma de ambos: λογ + ος = λογος (combinação)

iv. combinação = vocábulo digitado

2. Processam-se as combinações.

V. Estando no nominativo, registro 36:

1. decomposição da desinência do vocábulo de acordo com a desinência do banco:

a. vocábulo digitado: λογος

b. tabela → registro 36 → desinência do nominativo: ς (extensão desinencial (ED): 1)

c. decomposição do vocábulo (ED = 1):

i. extraindo-se o último caracter: λογος = λογο (provável raiz)

• comparação do ς (posição final) de λογος com ς (posição final) da desinência

• sendo iguais, efetua-se a soma

d. soma do que sobrou do vocábulo digitado (raiz) com a desinência do banco:

i. raiz, a partir da decomposição: λογ

ii. desinência do banco: ς

iii. soma de ambos: λογ + ς = λογος (combinação)

iv. combinação = vocábulo digitado

2. processam-se as combinações.

3. processam-se todos os registros em todos os casos.

Diante do processamento realizado, qual decisão deve ser tomada? Qual das combinações é a mais provável? Para tal, é necessário observar os indicadores fornecidos pelo programa, a saber:

1. ED – extensão desinencial – classificação decrescente;

2. % nos casos – busca no léxico ou no *corpus*;

3. existência, no vocabulário, da forma nominativa construída.

O alinhamento desses três indicadores, em vários contextos, aponta para a combinação mais provável. Vejamos:

Para o vocábulo **λογος**, após processamento, há as seguintes combinações**:**

1	1	2	90%	S	1		o	o
24	3	2	100%	S	1		jo	o
36	2	1	30%	S	1		ν	ν

Figura 3 – Combinações efetuadas pelo programa Nomes

3 - Nominativo Singular: **λογος**
Análise: nominativo relativo masc. singular: **λογος**
Tema em: **ο, λογο-**

λόγος-ου, ὁ, (Τ λόγο-), a palavra, o discurso, a razão

Figura 4 – Combinação 3

2 -Nominativo Singular: λογος
Análise: nominativo relativo interrogativo masc. e fem., relativo indefinido masc. e fem singular: λογος
Tema em: ν, λογον-

λόγος-ου, ὁ, (Τ λόγο-), a palavra, discurso, razão

Figura 5 – Combinação 2

1 - Nominativo Singular: λογος
Análise: nominativo nome masc. e fem., adjetivo masc., relativo masc. singular: λογος
Tema em: ο, λογο-

λόγος-ου, ὁ, (Τ λόγο-), a palavra, o discurso, a razão

Figura 6 – Combinação 1

A partir dos indicadores – extensão desinencial, porcentagem nos casos e existência, no vocabulário, da forma nominativa construída –, a combinação proposta pelo programa é a de número 3: ED = 2, 100% de ocorrência no *corpus* e a forma nominativa encontrada no vocabulário (S). Entretanto, o vocábulo λογος não é um relativo, como quer a combinação 3, mas um nome-substantivo, como mostra o vocabulário. Portanto, a combinação mais provável, imediatamente após a 3, é a de número 1.

VI. Tecnologias Adaptativas no Sistema Nominal do Grego Ático

Tendo em vista que dispositivos adaptativos podem ser utilizados em diversas situações, principalmente nas que pressupõem resoluções de problemas práticos e tomadas de decisão na solução de um problema, como possibilidades de aplicações da Tecnologia Adaptativa na área das Ciências Humanas, mais especificamente na Lingüística, destacamos aspectos da Adaptatividade que podem ser aplicados ao Sistema de Computação para o Estudo da Língua Grega Clássica.

Técnicas, métodos e ferramentas para Tecnologia Adaptativa bem como suas aplicações podem possibilitar que o sistema computacional, enquanto mecanismo subjacente, não-adaptativo, realize operações de automodificação em seu próprio comportamento – código –, em tempo de execução, de forma autônoma, sem interferência externa. Para tanto, é preciso estabelecer um conjunto de funções e regras adaptativas – camada adaptativa e programas adaptativos, respectivamente – para que seja possível alterar, a partir de determinados estímulos e ações, as configurações iniciais (de CF_1 para CF_2, de CF_2 para CF_3, de CF_3 para CF_n).

No nosso projeto de flexão nominal, a criação de regras adaptativas pode ser apropriada nos seguintes pontos:

a) reconhecimento automático de padrões do Sistema Nominal do Dialeto Ático a partir das pesquisas morfológicas feitas pelo Sistema ao Léxico exaustivo do TLG, com o objetivo de validar ou não as combinações dos casos da língua grega propostas pelo cálculo morfológico. A utilização dos recursos de aprendizagem de máquina através da construção de regras adaptativas pode levar à combinação morfológica correta, sem a necessidade de se recorrer ao léxico grego. Isso porque nosso sistema não indica, categoricamente, qual combinação nominal é a "verdadeira", mas propõe um conjunto de combinações morfológicas possíveis, reagrupando-as em ordem decrescente de ocorrências no léxico;

b) alargamento do universo lingüístico de forma a permitir que o dispositivo adaptativo estabeleça a flexão nominal não só para o grego, mas também para outras línguas sintéticas, a exemplo do hebraico, latim, sânscrito etc. – construção de um conjunto de regras para os comportamentos idênticos (desinências que não sofrem maiores alterações fonéticas, em função do que a raiz nominal não é afetada – maioria dos casos) e de um conjunto de regras para os comportamentos diferenciados (casos em que as alterações nas desinências acabam por afetar a raiz nominal). Acreditamos que isso possa ser formalizado e aplicado para todas as línguas sintéticas.

É dessa forma que vislumbramos resultados positivos de uma confluência do nosso sistema com a área da Inteligência Computacional – uma conexão do formalismo já desenvolvido a mecanismos adaptativos será eficiente para a geração de uma Base de Conhecimentos com os recursos complementares capazes de fornecer a combinação correta do grego ático e de permitir aplicação a outras línguas sintéticas.

VII. Conclusão

Enfatizamos, de um lado, que a pesquisa está em curso e, por isso, muitos aspectos podem sofrer alterações; de outro, que a idéia geral do projeto, a saber, a construção de um programa de computador que forneça um *conjunto de informações* sobre a flexão nominal do dialeto ático, está, a cada passo da nossa pesquisa, concretizando-se.

A orientação principal de se constituir um banco de dados para armazenar informações para tratamento computacional da língua grega já é uma realidade neste trabalho. Inúmeros benefícios poderiam ser listados, mas

ficamos apenas com o relacionado à eficácia e alta produção lingüísticas. A análise e o cálculo morfológicos efetuados pelo computador – que possiblita a flexão nominal (e futuramente a flexão verbal) – e a busca das formas nos textos dos autores, num espaço de tempo ínfimo – num *corpus de aproximadamente 99 milhões de palavras* –, são vantagens indiscutíveis do uso das Novas Tecnologias Digitais nas pesquisas lingüísticas. O emprego das Tecnologias Adaptativas descortina a possibilidade do reconhecimento automático de padrões na flexão nominal da língua grega e de outras línguas sintéticas.

REFERÊNCIAS

[1] BAILLY, Anatole. Dictionnaire grec français. 26e ed. Paris: Hachette, 196

[2] CAMLONG, André. Lexicométrie I - La correlation simple. O teste t de Student-Fisher ou a regressão linear simples? Toulouse: Université de Toulouse Le-Mirail, 2003.

[3] _____. Lexicométrie II - La corrélation simple et la corrélation multiple. Toulouse: Université de Toulouse Le-Mirail, 2003.

[4] _____. Lexicométrie III ou La lexicométrie avec Stablex. Toulouse: Université de Toulouse Le-Mirail, 2003.

[5] _____. Méthode d'analyse lexicale textuelle et discursive. Paris: C.R.I.C. & OPHRYS, 1996.

[6] _____. Stablex pratique. Toulouse: Teknea, 1991.

[7] _____; BELTRAN, Thierry. Stablex. Manuel d'utilisation. Toulouse: A.P.I, 1991.

[8] _____; _____; ZAPPAROLI, Zilda Maria. A propósito do Stablex - versão PC. São Paulo: Pirus Tecnologia, 2004.

[9] NETO, João José. Tecnologia Adaptativa. São Paulo: LTA - Laboratório de Linguagens e Tecnologias Adaptativas, set. 2004. Disponível em: http://www.pcs.usp.br/~lta/roteiro_estudo/index.html. Acesso em 6 de novembro de 2008.

[10] CANTÙ, Marcos. Dominando o Delphi 2. Trad. de Edimilson Kazwyoshi Miyasaki. São Paulo: Makron Books, 1996.

[11] CAVALCANTI, Edenis Gois. Construção de Software para o Estudo da Língua Grega Clássica: o Sistema Nominal do Dialeto Ático na Visão Temática. São Paulo, 2008. Tese (Doutorado em Letras - Lingüística) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo. /em desenvolvimento/

[12] MURACHCO, Henrique. Língua grega: visão semântica, lógica, orgânica e funcional. São Paulo: Discurso Editorial / Vozes, 2001.

[13] ZAPPAROLI, Zilda Maria; CAMLONG, André. Do léxico ao discurso pela Informática. São Paulo: EDUSP/FAPESP, 2002.