

# A BASE NACIONAL COMUM CURRICULAR E OS RESULTADOS DE SUA CONSULTA PÚBLICA: UMA LEITURA DO COMPONENTE CURRICULAR QUÍMICA

Ronaldo Spinelli Jr., Fernando L. Cássio  
*Universidade Federal do ABC*

**RESUMO:** Este trabalho apresenta uma leitura comparativa entre a primeira versão da Base Nacional Comum Curricular (BNCC) e a consulta pública *online* a este texto com a participação de professores de todo o Brasil. Realizou-se uma análise lexicográfica com o *software* IRaMuTeQ, que permitiu dividir o texto em classes de segmentos de texto (ST) com conteúdo semântico semelhante. Os resultados mostram que as contribuições dos professores possuem uma estrutura compartimentada. Ademais, os verbos utilizados permitem inferir, a partir dos princípios da taxonomia de Bloom, que os contextos de utilização dos verbos remetem a processos cognitivos de ordem mais baixa do que os encontrados no texto da BNCC.

**PALAVRAS-CHAVE:** Ensino de Química, Currículo, Análise Lexical.

**OBJETIVOS:** Apresentar uma leitura comparativa entre a primeira versão da Base Nacional Comum Curricular (BNCC) e as contribuições de professores de Química a este texto durante a consulta pública à BNCC, realizada pelo Ministério da Educação (MEC) entre 2015 e 2016. Essa “leitura” será realizada através de métodos de análise lexicográfica, com ênfase na análise do emprego dos verbos a partir de uma taxonomia de objetivos educacionais (taxonomia de Bloom).

## INTRODUÇÃO E MOTIVAÇÕES

As discussões públicas ao redor da Base Nacional Comum Curricular (BNCC) no Brasil têm gerado uma série de tensões, que embora inerentes ao espaço de disputas que forja as discussões de currículo, tem provocado debates acerbos na comunidade educacional do país.

A elaboração de referenciais curriculares nacionais é uma realidade há décadas em diversos países. A Coreia do Sul, onde o currículo nacional foi instituído na década de 1950, lançou a sétima versão de seu referencial curricular em 2000. Na Inglaterra, o currículo nacional foi instituído pela Lei da Reforma Educativa de 1988 (Brasil, 2014). Outros processos de implantação de currículos oficiais ensejaram debates públicos mais abertos, como na Austrália, onde uma grande parte das disputas se concentrou no risco da perda de autonomia das redes de ensino por uma envoltória potencialmente padronizadora e homogeneizante de documentos curriculares de alcance nacional.

Há um relativo consenso de que esse movimento de reformas curriculares é influenciado por mudanças na economia global, que afetam as relações sociais de trabalho. As reformas curriculares são, assim, uma forma de reorganizar as relações de poder entre diferentes grupos (Goodson, 2008).

Apesar das diferentes posições que se pode tomar neste debate, inclusive contrariamente à existência de uma base curricular com veleidades unificadoras e que enseje o desenvolvimento de mecanismos de controle para a sua legitimação simbólica (Apple, 1996), uma impressionante quantidade de dados foi gerada a partir da consulta pública realizada pelo Ministério da Educação (MEC) entre 2015 e 2016, e que agora constitui uma valiosa fonte primária para investigações sobre o currículo e suas representações. Em março de 2016, o MEC divulgou que o número absoluto de contribuições depositadas na plataforma da BNCC superara os 12 milhões, número incrivelmente alto se pensarmos que a maior parte das pesquisas sobre representações curriculares utilizam metodologias qualitativas e um número relativamente baixo de sujeitos de pesquisa.

## ANÁLISE LEXICAL E TAXONOMIA DE BLOOM

A análise de dados textuais com o auxílio de *softwares* de *text mining* permite lidar com grandes quantidades de texto, produzindo resultados de grande relevância aos estudos dos pensamentos, crenças e representações sociais: conteúdos simbólicos produzidos em relação a um determinado fenômeno, dados compostos basicamente pela linguagem (Camargo & Justo, 2013).

No caso de textos curriculares, pode-se interpretar os resultados das análises lexicográficas utilizando, por exemplo, uma taxonomia de objetivos educacionais como a de Bloom. Segundo Ferraz e Belhot (2010), Bloom e seus colaboradores definiram as características básicas dos domínios cognitivo, afetivo e psicomotor. O domínio cognitivo, relacionado ao ato de aprender ou dominar um novo conhecimento, é o mais utilizado. O domínio cognitivo está diretamente relacionado aos verbos utilizados da construção dos objetivos da aprendizagem. Na taxonomia de Bloom, essa classe de palavras se encontra agrupada em seis categorias: *conhecimento*, *compreensão*, *aplicação*, *análise*, *síntese* e *avaliação*, cada uma delas apresentando um nível específico de complexidade. Na categoria *conhecimento*, que apresenta uma menor complexidade, aparecem palavras como “definir”, “reconhecer”, “memorizar” e “recordar”, enquanto na categoria *aplicação*, de maior complexidade, são apresentados verbos como “demonstrar”, “descobrir” e “prever”. Já na categoria *avaliação*, a mais complexa, são apresentadas as palavras “comparar”, “criticar” e “decidir”.

## METODOLOGIA

### Fonte dos dados e consolidação do *corpus* textual

Foram produzidos dois *corpora* textuais: um a partir do texto original da primeira versão da BNCC (componente curricular: Química) e outro a partir das contribuições dos professores brasileiros coletadas no período de consulta pública ao texto (setembro de 2015 a março de 2016), publicamente disponíveis em plataforma do MEC.

Os textos deram origem a dois arquivos de texto (codificação UTF-8) acrescidos de metadados, como a região e o estado do Brasil de onde provieram as contribuições, o ano escolar a que se refere a contribuição ou o objetivo de aprendizagem (1º, 2º ou 3º do Ensino Médio), entre outros. Após a leitura integral do material, foi realizada a hifenização do *corpus*, de modo que expressões como “Tranformações\_Químicas” ou “Ensino\_Médio” não fossem reconhecidas pelo *software* como palavras separadas.

O *corpus* das contribuições apresentou 28.429 palavras, em um total de 600 contribuições, referentes ao primeiro (N = 209), segundo (N = 181) e terceiro ano do Ensino Médio (N = 210); enquanto o *corpus* da BNCC apresentou 3.176 palavras, em um total de 54 textos, 18 textos referentes ao primeiro, 20 textos ao segundo e 16 textos ao terceiro ano do Ensino Médio.

Foram realizadas análises preliminares para encontrar os parâmetros que maximizassem os segmentos de textos analisados evitando que informações relevantes fossem descartadas durante o processo de análise feita pelo *software*.

### Análise de dados verbais

As análises foram realizadas utilizando o *software* IRaMuTeQ (*Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*), que é gratuito e ancorado no ambiente de computação estatística R. O IRaMuTeQ possui diferentes rotinas para análises lexicográficas (quantidade e frequência de palavras) e sua visualização (nuvens de palavras e diagramas de Pareto), classificação hierárquica descendente (CHD) e análises de especificidades e similitude. Neste estudo foram geradas análises de CHD para os dois *corpora*, cujos resultados foram comparados.

O método da CHD (Reinert, 1987) classifica os segmentos de texto (ST) do *corpus* em função de seus respectivos léxicos, sendo o conjunto desses segmentos repartido com base na frequência das formas reduzidas (raízes das palavras). Dessa forma são obtidas unidades de contexto elementar (UCE) com um vocabulário semelhante entre si, mas diferente das UCE de outras classes (Camargo & Justo, 2013). O *software* permite visualizar as associações entre as palavras através de dendrogramas e de diagramas de análise fatorial de correspondência (AFC) (Lebart *et al.*, 1998).

As análises dos dois *corpora* textuais foram realizadas em duas etapas: na primeira foram incluídos somente substantivos e adjetivos; na segunda foram incluídos somente os verbos. Textos curriculares são fortemente marcados pelo uso de verbos no infinitivo, daí a análise das classes lexicais a partir dos verbos que as caracterizam ser uma forma interessante de “ler” os *corpora* textuais a partir dos objetivos educacionais apresentados. Uma análise preliminar serviu para otimizar o parâmetro NCT (número de classes terminais) que maximizam o percentual de segmentos de texto classificados (STC) (Tabela 1).

Tabela 1.  
Números máximos de classes  
terminais (NCT) utilizados nas análises.

ANÁLISE	NCT
BNCC	8
BNCC – somente verbos	8
Consulta pública	8
Consulta pública – somente verbos	14

Esse parâmetro define o número máximo de classes em que o *software* poderá agrupar as palavras analisadas, em geral o número de classes criadas é menor que o valor do parâmetro adotado (o valor *default* do IRaMuTeQ é 10). Lahlou (2012), em uma crítica bastante fundamentada sobre o uso de métodos de *text mining* em pesquisas sobre representações sociais, afirma que o *software* é apenas um instrumento e não o método, sendo necessário que o analista busque maneiras de usá-lo da melhor

forma possível. Com exceção da análise do *corpus* da BNCC só com verbos, em que só foram classificados 45,92% dos ST (o que se justifica pelo tamanho reduzido deste *corpus*), as CHD foram obtidas STC sempre superiores a 75% (menor quantidade de ST que não fazem parte de nenhuma classe).

## RESULTADOS

Na tabela 2 são apresentados alguns dos dados de saída da CHD para as quatro análises realizadas:

Tabela 2.  
Dados das análises dos dois *corpora* textuais.

	BNCC	CONSULTA PÚBLICA	BNCC (SOMENTE VERBOS)	CONSULTA PÚBLICA (SOMENTE VERBOS)
	Figura 1 ( esq.)	Figura 1 ( dir.)	Figura 2 ( esq.)	Figura 2 ( dir.)
N	54	600	54	600
ST	98	996	98	996
STC	84 (85,71%)	804 (80,72%)	45 (45,92%)	819 (82,23%)
Classes	5	3	3	5

N: número de textos analisados; ST: número de segmentos de texto; STC: segmentos de texto classificados (%)

A seguir, na Figura 1, são apresentados os diagramas de análise fatorial de correspondência (AFC) para a BNCC e para o conjunto de contribuições dos professores coletado durante a consulta pública:

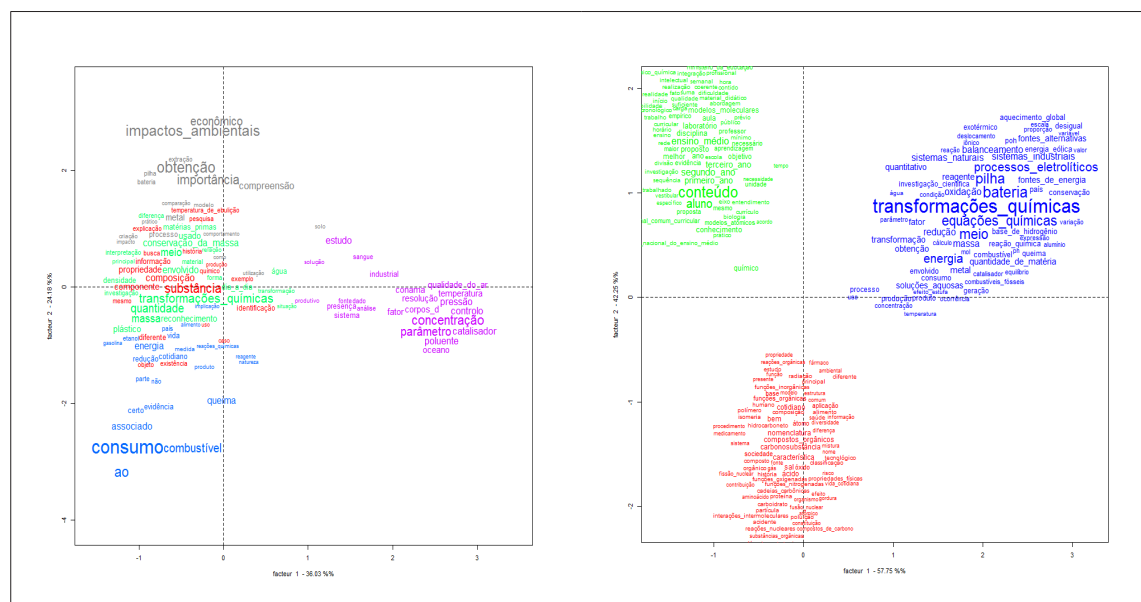


Fig. 1. Diagramas de AFC para a primeira versão da BNCC (à esquerda) e para a consulta pública (à direita). O tamanho da fonte é proporcional ao grau de associação das palavras às respectivas classes ( $\chi^2$ ).

Embora o tamanho da fonte nos gráficos não seja muito convidativo à leitura das palavras classificadas, os diagramas dão a ver diferenças importantes em relação à estrutura geral dos dois “textos”. Há uma diferença significativa na forma como as classes estão concentradas nos dois diagramas. No diagrama referente à BNCC, é possível perceber uma correlação forte entre 4 das 5 classes apresentadas. No diagrama das contribuições, ao contrário, observa-se uma maior homogeneidade no conteúdo semântico das três classes, que aparecem bastante isoladas umas das outras e fortemente vinculadas a ST que remetem a conteúdos de Química.

Esse resultado indica que as palavras utilizadas no texto da BNCC estão associadas a mais de uma classe, o que sugere uma certa transversalidade na estrutura do texto. Este comentário, no entanto, não constitui juízo de valor sobre o seu conteúdo.

Já o diagrama relativo à consulta pública mostra uma estrutura bastante compartimentada, com uma classe bastante vinculada a temas sobre “transformações\_ químicas” (representada em azul) e reações em solução, outra a assuntos sobre química descritiva (incluindo, por exemplo, “química\_ orgânica”, em vermelho) e uma terceira classe com ST relacionados ao trabalho didático-pedagógico da Química (representada em verde).

Na Figura 2 são reapresentados os diagramas de AFC para os dois *corpora*, agora considerando apenas os verbos.

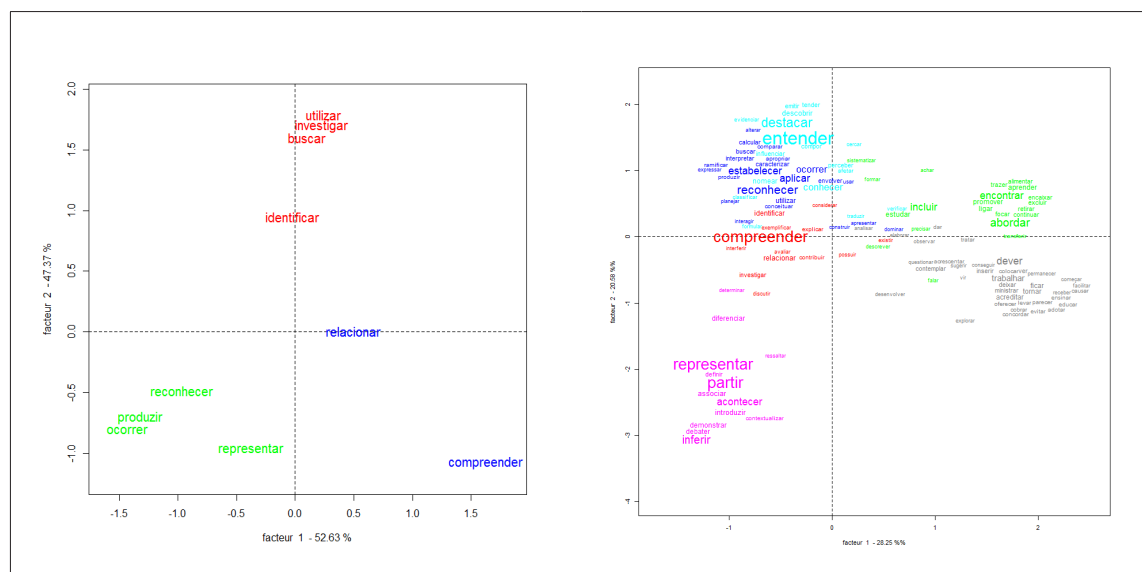


Fig. 2. Diagramas de AFC para a primeira versão da BNCC (à esquerda) e para a consulta pública (à direita), incluindo somente os verbos. O tamanho da fonte é proporcional ao grau de associação das palavras às respectivas classes ( $\chi^2$ ).

A diferença na quantidade de palavras representadas nos dois diagramas pode ser justificada pela diferença de tamanho entre os dois *corpora*: o *corpus* da consulta pública contém 28.429 palavras, enquanto o *corpus* da BNCC conta com 3.176 palavras.

A partir da associação desses verbos entre si e com as diferentes classes (Figura 2, especialmente a da direita) e recorrendo à taxonomia de Bloom, podemos identificar os domínios cognitivos, nas contribuições feitas pelos docentes aparecem em destaque os verbos “compreender”, “reconhecer”, “inferir” e “representar”. Segundo a taxonomia de Bloom, estes verbos se referem a operações mais simples no domínio cognitivo, mais especificamente encontram-se na capacidade de lembrar informações e conteúdos e de compreender e dar significado aos conteúdos, no diagrama também aparecem palavras que

sugerem uma maior complexidade palavras como “estabelecer” e “diferenciar” entretanto não são tão comuns no texto. (Ferraz & Belhot, 2010)

No texto da BNCC a frequência e a variedade de verbos é menor do que nos resultados da consulta pública, o que condiz com o tamanho do *corpus*. Ali também aparecem verbos relacionados a processos cognitivos mais simples como “relacionar”, “identificar”, “reconhecer” e “compreender”, mas também verbos associados a processos cognitivos de ordem mais alta, como “investigar” e “relacionar”, estando este último associado à categoria onde se pretende desenvolver a habilidade de agregar partes para formar um todo, sendo necessário um nível de abstração muito maior (Ferraz & Belhot, 2010). Outro aspecto interessante é o fato de não haver um verbo predominante no texto da BNCC: todos aparecem de forma equivalente, pois o tamanho das palavras no gráfico é praticamente o mesmo.

## CONCLUSÕES

Foi possível verificar que as contribuições dos professores na consulta pública sobre a primeira versão da BNCC possuem uma estrutura fortemente compartimentada, que contrasta com uma aparente transversalidade do texto analisado. Ademais, os verbos utilizados nos dois *corpora* permitem inferir, a partir dos princípios da taxonomia de Bloom, que embora o número de palavras seja maior no *corpus* textual da consulta pública, os contextos de utilização dos verbos remetem a processos cognitivos de ordem mais baixa do que os encontrados no texto-proposta do futuro currículo oficial do Brasil.

## REFERÊNCIAS BIBLIOGRÁFICAS

- APPLE, M.W. (1996). *El conocimiento oficial: La educación democrática en una era conservadora*. Barcelona: Paidós.
- BRASIL, & MINISTÉRIO DAS RELAÇÕES EXTERIORES. (2014). *Mundo afora: Educação Básica e Ensino Médio*. Brasília, DF: Ministério das Relações Exteriores.
- CAMARGO, B.V., & JUSTO, A.M. (2013). IRAMUTEQ: Um software gratuito para análise de dados textuais. *Temas em Psicologia*, 21(2), 513-518.
- FERRAZ, A.P.C.M., & BELHOT, R.V. (2010). Taxonomia de Bloom: Revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. *Gest. Prod.*, 17(2), 421-431.
- GOODSON, I. F. (2008). *As políticas de currículo e de escolarização*. Petrópolis, RJ: Vozes.
- LAHLOU, S. (2012). Text mining methods: An answer to Chartier and Meunier. *Papers on Social Representations*, 20(38), 1-7.
- REINERT, M. (1987). Classification descendante hiérarchique et analyse lexicale par contexte: Application au corpus des poésies d'A. Rimbaud. *Bulletin de Méthodologie Sociologique*, 13, 53-90.