

Análise Textual Discursiva apoiado por software: IRaMuTeQ e a análise de subcorpus

Software-supported Discursive Textual Analysis: IRaMuTeQ and subcorpus analysis

Valderez Marina do Rosário Lima ¹ <https://orcid.org/0000-0002-2676-5840>

Marcelo Prado Amaral-Rosa ² <https://orcid.org/0000-0002-3294-8141>

Maurivan Güntzel Ramos ³ <https://orcid.org/0000-0002-2586-0723>

¹Escola Politécnica, Programa de Pós-Graduação em Educação em Ciências e Matemática e Escola de Humanidades, Programa de Pós-Graduação em Educação da PUCRS – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre/RS, Brasil

^{2,3} Escola Politécnica, Programa de Pós-Graduação em Educação em Ciências e Matemática da PUCRS – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre/RS, Brasil

Resumo:

Introdução: o foco deste trabalho é o processo de geração de categorias a partir da análise de subcorpus de dados qualitativos que são apoiadas no software IRaMuTeQ. A relevância se dá pela ausência de análises de subcorpus na literatura, inclusive de diferentes áreas, na utilização do referido software. O tema central são as nuances relacionadas aos Clubes de Ciências em escolas de Educação Básica. Objetivo: o objetivo foi apresentar as análises de subcorpus no procedimento de categorização que utilizam o método da Análise Textual Discursiva apoiado pelo software IRaMuTeQ, com vistas à compreensão sobre a geração das categorias intermediárias. Método: É uma pesquisa de abordagem qualitativa. O corpus e os subcorpus de análises estão baseados em quatro diálogos sobre o mundo dos Clubes de Ciências entre duas pesquisadoras experientes na temática. O método analítico adotado é a Análise Textual Discursiva apoiado pelo software IRaMuTeQ. No software, tanto para tratamento de corpus quanto de subcorpus, a análise de base foi a Classificação Hierárquica Descendente. Resultados: identificou-se a emergência de cinco categorias intermediárias, sendo possível observar, enquanto principal resultado, a partir das análises de subcorpus, o reconhecimento da composição de cada categoria intermediária por outras cinco a sete categorias iniciais. Conclusões: o procedimento de análise de subcorpus é de suma importância para a compreensão das construções de cada categoria intermediária emergente no software adotado para apoio ao método da Análise Textual Discursiva, configurando-se, enquanto um nível avançado, com agrupamentos, da etapa inicial da categorização, não sendo propriamente o primeiro movimento procedimental, mas ainda dentro da etapa inicial de categorização.

Palavras-chave: Análise Textual Discursiva; Análise Qualitativa; IRaMuTeQ; Subcorpus.

Abstract:

Introduction: the focus of this work is the process of generating categories from the analysis of subcorpus of qualitative data that are supported in the IRaMuTeQ software. The relevance is due to the absence of subcorpus analysis in the literature, including from different areas, in the use of said software. The central theme is the nuances related to Science Clubs in basic education schools. Objective: the aim was to present subcorpus analyses in the categorization procedure that use the method of textual discursive analysis supported by the IRaMuTeQ software, with a view to understanding the generation of intermediate categories. Method: it is a qualitative approach research. The corpus and subcorpus of analysis is a dialogue about the world of Science Clubs between two researchers experienced in the subject. The analytical method adopted is the Textual discursive analysis supported by the IRaMuTeQ software. In the software, both for Corpus and subcorpus analysis, the basic analysis was the descending hierarchical classification. Results: the data analysis identified the emergence of five intermediate categories, and it is possible to identify as the main result, from subcorpus analyses, the identification of the composition of each intermediate category by another five to seven initial categories. Conclusions: the subcorpus analysis procedure is of paramount importance for the understanding of the constructions of each intermediate category emerged in the software adopted to support the method of textual discursive analysis, being configured as an advanced level, with groupings, of the initial categorization stage, not being the first movement, but still within the initial stage.

Keywords: Discursive Textual Analysis; Qualitative Analysis; IRaMuTeQ; Subcorpus.

Submissão: 18/03/2021

Aceitação: 28/04/2021

1. Introdução

O ponto nevrálgico deste texto são as nuances relacionadas aos Clubes de Ciências (Mancuso, Lima & Bandeira, 1996; Viêra & Lima, 2016; De Prá & Tomio, 2014, Camors, 2013; Bernet, 2013; Lima, 1998; Albuquerque, Lima & Rosito, 2016) em escolas da Educação Básica a partir do ponto de vista analítico e das vivências de pesquisadoras experientes frente à temática e formação de professores no estado do Rio Grande do Sul, Brasil (Rosito & Lima, 2020). Há cerca de 50 anos, buscam-se possibilidades para a forma convencional de ensino de Ciências na América do Sul (Camors, 2013) e os Clubes de Ciências são uma alternativa, pois se configuram enquanto espaços organizados em

torno de interesses coletivos (Tomio & Hermann, 2019) sobre temáticas, em geral, paralelas aos conteúdos programáticos constantes nas ementas das disciplinas escolares formais.

Diante do exposto, com vistas ao entendimento em detalhes sobre o tema em questão, adotou-se, como método analítico, a Análise Textual Discursiva (Moraes & Galiuzzi, 2016) apoiado pelo software IRaMuTeQ (Ratinaud, 2014; Camargo & Justo, 2013; Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020). A ênfase da Análise Textual Discursiva é concentrada nas possibilidades de dados de subcorpus gerados a partir do software de apoio IRaMuTeQ (Moraes & Galiuzzi, 2016; Lima & Ramos, 2017; Moraes, Galiuzzi & Ramos, 2013; Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020), uma vez que softwares destinados à pesquisa com dados qualitativos estão cada vez mais presentes e necessários devido à gama de informações geradas a serem gerenciadas (Gray, 2012; Stake, 2011; Mayring, 2014).

A abordagem do método em questão é de natureza qualitativa (Moraes & Galiuzzi, 2016). Convém salientar, para evitar entendimentos errôneos, dois aspectos referentes ao apoio do software IRaMuTeQ: i) mantêm-se a abordagem qualitativa do método inalterada (Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020), devido ao seu forte caráter interpretativo hermenêutico (Moraes & Galiuzzi, 2016).; e ii) o papel do pesquisador não é substituído pelo do software, pois é insuficiente a aplicação de procedimentos operacionais sem o requerido entendimento sobre aquilo que se está a fazer frente às análises dentro do software (Costa & Amado, 2018).

A principal contribuição da Análise Textual Discursiva é a produção de novas compreensões sobre o material em foco (Moraes & Galiuzzi, 2016; Lima & Ramos, 2017; Moraes, Galiuzzi & Ramos, 2013). A organização está baseada em três etapas: i) identificação de *unidades de sentido* (unitarização); ii) estabelecimento de aglutinações em cadeia (categorização); e iii) construção de compreensões inéditas (metatextos) (Moraes & Galiuzzi, 2011; Lima & Ramos, 2017; Moraes, Galiuzzi & Ramos, 2013). Já entre os benefícios do IRaMuTeQ estão a gratuidade do software, execução de análises de dados textuais em diferentes níveis e a agilidade frente ao tratamento e geração de informações (Camargo & Justo, 2013; Kami et al., 2016; Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020). Neste estudo, a atenção está concentrada na análise de subcorpus decorrente da análise de Classificação Hierárquica Descendente (CHD) (Ratinaud, 2014).

A relevância do estudo é declarada com base em três aspectos, a saber: i) a ausência do tratamento de subcorpus em análises qualitativas apoiadas no software IRaMuTeQ, independentemente do método de análise adotado; ii) entendimento sobre a emergência das categorias intermediárias, sendo esse um ponto nebuloso na utilização de softwares qualitativos; e iii) apoio do IRaMuTeQ no nível da categorização inicial do método da Análise Textual Discursiva, até então, ponto inexplorado da relação entre ambos. Com isso, declara-se o ineditismo das perspectivas analíticas para a pesquisa qualitativa na utilização do IRaMuTeQ enquanto ferramenta de apoio ao método da Análise Textual Discursiva (Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020; Moraes & Galiuzzi, 2016; Lima & Ramos, 2017; Moraes, Galiuzzi & Ramos, 2013).

Assim sendo, a questão norteadora é: *De que modo as análises de subcorpus no software IRaMuTeQ auxiliam a compreensão sobre as formações das categorias intermediárias na utilização do método de Análise Textual Discursiva?*

Nesse sentido, o objetivo primordial foi apresentar a importância das análises de subcorpus no procedimento de categorização que utilizam o método da Análise Textual Discursiva apoiado pelo software IRaMuTeQ, com vistas à compreensão detalhada sobre a geração das categorias intermediárias.

Com relação à estrutura, apresentam-se três seções: i) *procedimentos metodológicos*, com foco em expor as estratégias assumidas; ii) *resultados e discussões*, cujo cerne são os resultados alcançados com as análises de subcorpus do material em análise no software IRAMUTEQ e as conexões com a literatura especializada; e iii) *conclusões*, em que se retoma a questão norteadora e se declara as contribuições das análises de subcorpus no software IRaMuTeQ no processo de categorização do método da Análise Textual Discursiva.

2. Procedimentos Metodológicos

2.1 Objeto de Estudo

O objeto de estudo são contribuições das análises de subcorpus para o processo de categorização do método de Análise Textual Discursiva (Moraes & Galiuzzi, 2016; Lima & Ramos, 2017; Moraes, Galiuzzi & Ramos, 2013) apoiado pelo software IRaMuTeQ (Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020). Sublinha-se que a análise de subcorpus é possível a partir da execução da análise de *Classificação Hierárquica Descendente* (CHD) sobre o corpus devidamente preparado e inserido no software IRaMuTeQ (Ratinaud, 2014; Camargo & Justo, 2013; Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020). A temática de análise são as nuances que englobam as essências, organização, aprendizagens docentes e atividades investigativas de Clubes de Ciências nas escolas de Educação Básica no Brasil a partir de diálogos acadêmicos entre pesquisadoras com vivências e experiências sobre o assunto.

2.2 Participantes e Corpus de Análise

Os participantes são duas pesquisadoras da área de Educação em Ciências com larga experiência na docência em nível de Ensino Secundário e Ensino Superior. As informações a serem analisadas são baseadas em quatro diálogos, fundamentados na experiência e no fazer docentes de ambas, com distintos pontos norteadores, a saber: i) *as essências de um Clube de Ciências*; ii) *a organização de um Clube de Ciências*; iii) *as aprendizagens docentes no Clube de Ciências*; e iv) *as atividades investigativas desenvolvidas em um Clube de Ciências*. A principal intenção dos quatro diálogos é a divulgação de ideias renovadas sobre os temas que circundam o ensino de Ciências por meio dos Clubes de Ciências (Rosito & Lima, 2020). Vale frisar que os pontos norteadores não se configuram, nesta situação, enquanto categorias *a priori*, uma vez que não eram limitadores de tópicos, e sim funcionaram como um início para os diálogos.

Os diálogos ocorreram ao longo do ano de 2019. Foram cinco encontros com duração aproximada de 120 min cada. Surgiram em decorrência do desejo em atualizar informações sobre os Clubes de Ciências e culminou na confecção de um livro intitulado

Conversas sobre Clubes de Ciências (Ibid.). Acentua-se que a realização da análise sobre o referido material vem ao encontro de tal intenção, uma vez que a busca pelo entendimento das nuances do material, pode gerar desdobramentos analíticos e teóricos em projetos futuros.

No software IRaMuTeQ, o corpus de análise contemplou 16.316 palavras ao total, com 61 narrativas individuais (Textos), distribuídos em equilíbrio entre ambas pesquisadoras (n=32; n=29), e média de ~268 palavras por texto e média de ~35 por Segmentos de Textos (STs). O total de Segmentos de Textos formados foi de 460 e o índice de retenção/aproveitamento para o corpus foi de ~83% (376 STs). O total de palavras de ocorrência única foi de 799, correspondendo ao índice de Hapax 4.90%.

No subcorpus, total de Segmentos de Textos formados nos subcorpus foi de 352 e o índice de retenção/aproveitamento médio foi de ~73%, com o índice médio de Hapax de ~14% das ocorrências.

Com relação às análises realizadas, para o corpus, a análise base foi a *Classificação Hierárquica Descendente (CHD)*. Ela é necessária para a geração das categorias intermediárias. Para a geração de subcorpus, adotou-se, como parâmetro, os 100 Segmentos de Textos com maior densidade (score) em cada categoria intermediária, pois esse foi o número máximo alcançado nos testes de geração de cada subcorpus. E as análises utilizadas foram, de mesma forma que para o corpus, a *Classificação Hierárquica Descendente (CHD)*.

Com relação às decisões para a entrada dos dados no software IRaMuTeQ (Ratinaud, 2014), adotou-se o seguinte: i) idioma: *português*; ii) construção de STs: *ocorrências*; iii) tamanho de STs: *40*; iv) geração de subcorpus: *pontuação de score Absoluta*; e v) número de STs: *100*. Os demais pontos de entrada são os padrões pré-selecionados do próprio software IRaMuTeQ (Ibid.).

2.3 Análise dos Dados

A análise dos dados teve por base os preceitos da Análise Textual Discursiva (Moraes & Galiuzzi, 2016) apoiado pelo software IRaMuTeQ (Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020). Diante das três etapas procedimentais da Análise Textual Discursiva (unitarização, categorização e produção de metatextos), a ênfase aqui é apenas sobre a etapa de categorização (Moraes & Galiuzzi, 2016). A decisão é pautada nos três pontos a seguir: i) o processo de unitarização, via software IRaMuTeQ, ocorre de modo concomitante ao processo de categorização com a geração de STs, sendo apenas velada a sua apresentação visual; ii) a construção de metatextos demanda espaço e atenção que é inviável; e iii) o cerne aqui são as contribuições das análises de subcorpus, o que demanda apresentações de dados peculiares.

As análises realizadas para corpus e subcorpus são dois procedimentos distintos. Porém, são processos dependentes e sequenciais, tendo a mesma base de dados enquanto alicerce de informações. Portanto, para as análises de subcorpus, realizou-se: i) a análise do corpus frente à *Classificação Hierárquica Descendente (CHD)* (Ratinaud, 2014; Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020); e ii) análise de CHD para o subcorpus de cada uma das categorias intermediárias geradas dos 100 STs de maior densidade. A maneira realizada é a única alternativa possível para análise de subcorpus,

sem manipulação e preparação secundária das informações de modo prévio ao procedimento de inserção no software IRaMuTeQ.

Desse modo, ratifica-se que as análises interpretativas sobre as ramificações decorrentes do corpus, na análise de CHD, não se configuram em análise de subcorpus. Isso se deve ao fato de que é preciso gerar dados, em primeira instância, sobre o corpus para posterior geração de análise de subcorpus de cada categoria intermediária, de modo individualizado. Endossa-se ainda que, tanto a análise de corpus quanto de subcorpus, não são análises mecanizadas ou automatizadas que dispensam a interpretação do pesquisador sobre os dados originados, estando assim alinhado ao preconizado no método da Análise Textual Discursiva (Moraes & Galiazzi, 2016; Lima & Ramos, 2017; Moraes, Galiazzi & Ramos, 2013).

3. Resultados e Discussões

Para obtenção dos resultados a partir do método de Análise Textual Discursiva (Moraes & Galiazzi, 2016) apoiado pelo software IRaMuTeQ (Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020), tanto referente ao corpus quanto ao subcorpus das informações acerca das nuances dos Clubes de Ciências, utilizaram-se quatro diálogos entre duas pesquisadoras experientes da área do ensino de Ciências (Rosito & Lima, 2020).

Convém lembrar que o foco recai sobre o processo de categorização, e que o procedimento de unitarização, via software IRaMuTeQ, ocorre de modo concomitante e automatizado, baseado em cálculos estatísticos, sem a interferência do pesquisador (Martins et al., 2020).

Um dado importante para entrevistas, narrativas e diálogos na pesquisa qualitativa é a saturação (Gray, 2012; Stake, 2011). O software IRaMuTeQ fornece o coeficiente de Hapax, na execução da análise estatística básica, para todas as informações inseridas para análise (Ratinaud, 2014), e pode ser utilizado como indício da ocorrência [ou não] da saturação do corpus e também de subcorpus (Martins et al., 2020). No corpus, apenas 4.90% das ocorrências foram únicas, sugerindo que os quatro diálogos ocorreram em convergência com os pontos norteadores propostos (ver subseção 2.2), sem perda de foco ou divagações nos discursos, aspecto que demonstra o rigor na condução enquanto método de recolha de dados dos mesmos.

Já referente ao subcorpus, a média do coeficiente de Hapax foi de 14,41% (Δ 9.94% a 21,47%) para as cinco categorias intermediárias geradas. Com relação aos valores de Hapax, sublinha-se que quanto maior percentual, mais termos únicos estão presentes nos discursos, o que pode indicar tópicos inéditos da interação. Todavia, aqui, a elevação percentual de Hapax percebida frente ao subcorpus decorre do transcorrer natural dos diálogos, com alterações de momentos de domínio da narrativa entre os interlocutores, fazendo com que determinados termos se sobressaíam sobre outros, mas que não configura um novo rumo sobre o tema no discurso analisado. Assim, considera-se aceitável e compreensível a diferença apresentada entre os coeficientes de Hapax, para esse contexto analisado. Contudo, pondera-se que ainda são necessários mais estudos da relação entre o coeficiente de Hapax e saturação de discursos com outros *corpora* para melhor potencial de aproveitamento e compreensão do indicador em análises na pesquisa qualitativa.

Frente ao processo de categorização na Análise Textual Discursiva (Moraes & Galiuzzi, 2016; Lima & Ramos, 2017; Moraes, Galiuzzi & Ramos, 2013) apoiado pelo software IRaMuTeQ (Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020), executou-se a análise de *Classificação Hierárquica Descendente (CHD)* (Ratinaud, 2014). Na análise de CHD (Fig. 1), foi gerado um dendrograma com cinco *classes*, nomenclatura padrão do IRaMuTeQ, as quais são consideradas *categorias intermediárias* no método da Análise Textual Discursiva (Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020). O sentido da leitura é da esquerda para direita e o dendrograma está dividido em níveis de ramificações (R): i) R1, somente a categoria 5; e ii) R2, subdividida em R2(1) (categoria 3); e R2(2), com subramificações R2(2a) (categoria 1) e R2(2b) (categoria 4). Todos dendrogramas apresentados seguem a mesma lógica de indicações para as ramificações.

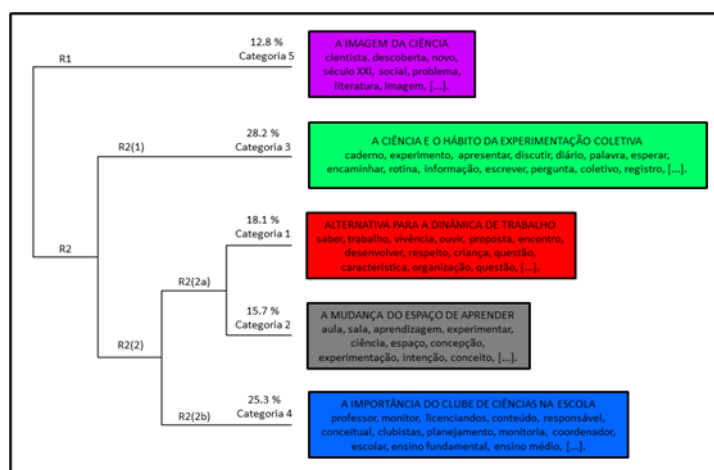


Figura 1: Categorias intermediárias acerca das nuances dos Clubes de Ciências.

Fonte: Dados do IRaMuTeQ adaptados.

As categorias intermediárias são formadas por aproximação e distanciamentos das Unidades de Sentido (Moraes & Galiuzzi, 2016) que se dá pela formação de Segmentos de Textos (Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020) de acordo com as frequências das ocorrências lematizadas de seus vocabulários, até chegar na estrutura mais estável (Ratinaud, 2014; Camargo & Justo, 2013; Veraszto et al., 2018). Os títulos das cinco categorias intermediárias são frutos interpretativos dos pesquisadores frente às unidades de sentido (Moraes & Galiuzzi, 2016) de maior densidade (*score*) em cada categoria intermediária. Dessa forma, as categorias intermediárias, suas cores e respectivas concentrações são: i) *alternativa para a dinâmica de trabalho* (vermelho – 18.1%); ii) *a mudança do espaço de aprender* (cinza – 15.7%); iii) *a ciência e o hábito da experimentação coletiva* (verde – 28.2%); iv) *a importância do Clube de Ciências na escola* (azul – 25.3%); e v) *a imagem da Ciência* (roxo – 12.8%). As concentrações totais somam ~83%, pois essa é o índice de retenção aproveitado do corpus. Vale lembrar que a retenção preconizada é de 70-75% (Ratinaud, 2014; Camargo & Justo, 2013).

Com suporte na CHD do corpus, executou-se as análises de subcorpus. No IRaMuTeQ, as análises de subcorpus, são realizadas para cada categoria intermediária, sendo essa a única opção de análise (Ratinaud, 2014). A intenção, com relação ao método da Análise Textual Discursiva, é a busca de subsídios para a compreensão frente às estruturas de formação de cada categoria intermediária (Moraes & Galiuzzi, 2016). Convém ressaltar

que esse tipo de análise é um ponto de escuridão na literatura, quando o assunto é o uso do IRaMuTeQ, uma vez que não são abordadas.

Para as análises de subcorpus, baseadas nos 100 Segmentos de Textos de maior densidade de categoria intermediária, executou-se de mesmo modo que para o corpus as análises de CHD. Na análise de CHD, o material gerado é sempre um dendrograma. A diferença agora que são referentes às categorias iniciais no método da Análise Textual Discursiva (Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020). Em função da disponibilidade de espaço, aqui é apresentado apenas a imagem global das categorias intermediárias e suas respectivas categorias iniciais (Fig.2) e, logo em seguida, é abordado, em mais detalhes, apenas a construção das categorias iniciais referentes à categoria 1 – *Alternativa para a dinâmica de trabalho* (vermelho) (Fig.3).

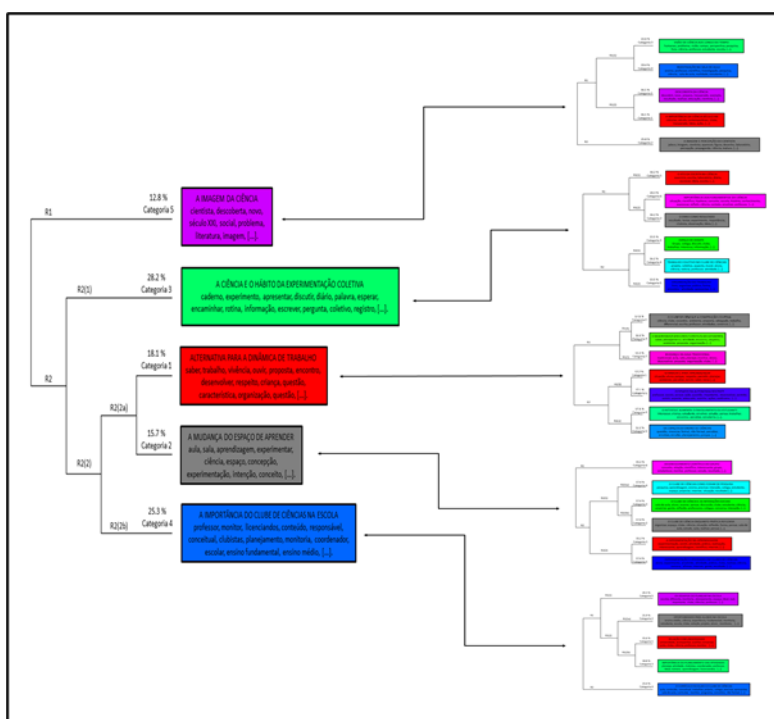


Figura 2: CHD da categorização global – categorias intermediárias e categorias iniciais.

Fonte: Dados do IRaMuTeQ adaptados.

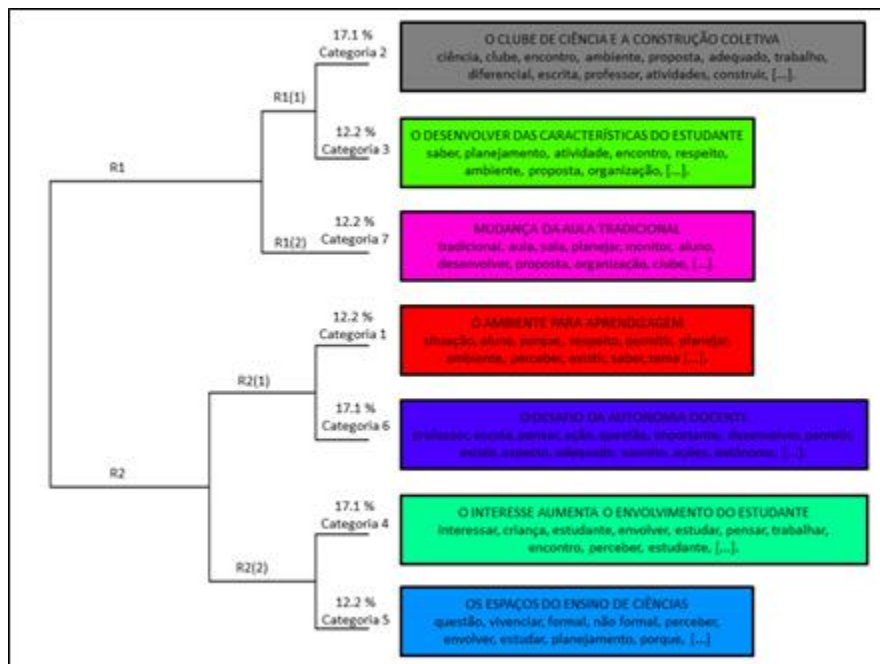


Figura 3: CHD das categorias iniciais – categoria interm. 1 - *Alternativa para a dinâmica de trabalho.*

Fonte: Dados do IRaMuTeQ adaptados.

De acordo com as frequências das ocorrências lematizadas dos STs que compõem a categoria 1 – *Alternativa para a dinâmica de trabalho* (Fig.1, vermelho), a estrutura mais estável (Ratinaud, 2014; Camargo & Justo, 2013; Veraszto et al., 2018) apresentou retenção de 82% e sete categorias iniciais, agrupadas em três ramificações internas: i) em R1, as categorias iniciais 2, 3 e 7; ii) em R2(1), as categoriais iniciais 1 e 6; e iii) em R2(2), as categoriais iniciais 4 e 5. Para melhor entendimento da relação entre as categorias iniciais, faz-se necessário as análises fatoriais de correspondência (AFC), contudo, aqui não são apresentadas devido ao espaço disponível diante das necessárias demonstrações de dados e argumentações.

Os títulos, assim como nas categorias intermediárias, também são decorrentes da interpretação dos pesquisadores com base nos Segmentos de Textos de maiores scores para cada categoria inicial. No caso, a formação da categoria intermediária 1, em questão, teve 41 STs totais entre as sete categorias iniciais formadas, o que corresponde a ~12% do total de 352 STs aproveitados para as análises de subcorpus. Por fim, anuncia-se que a realização das três etapas da Análise Textual Discursiva (Moraes & Galiazzi, 2016) apoiada por software (Ramos, Lima & Amaral-Rosa, 2019; Martins et al., 2020) exige fôlego. Salienta-se que a intenção foi avançar um pouco mais sobre o processo de categorização do referido método, tendo como ponto central de atenção à formação das intermediárias, por meio das categoriais iniciais com a análise de subcorpus no software voltado para análises qualitativas IRaMuTeQ (Ratinaud, 2014).

4. Conclusões

O princípio deste trabalho foi apresentar as análises de subcorpus no procedimento de categorização que utilizam o método da Análise Textual Discursiva apoiado pelo software IRaMuTeQ, com vistas à compreensão sobre a geração das categorias intermediárias. Assim, buscou-se responder à questão norteadora: *De que modo as*

análises de subcorpus no software IRaMuTeQ auxiliam a compreensão sobre as formações das categorias intermediárias na utilização do método de Análise Textual Discursiva?

A considerar as análises de dados, expõem-se sobre as análises de subcorpus:

- i) configura-se enquanto movimentos internos, sequenciais do corpus e individuais para cada categoria intermediária, não sendo possível a execução de outra maneira;
- ii) essencial para compreensão sobre as formações das categorias intermediárias, pois, na ausência, a primeira etapa do processo de categorização não é atingida;
- iii) caracteriza-se enquanto um nível avançado da categorização inicial, com agrupamentos (decorrentes da etapa de unitarização), não sendo propriamente o primeiro movimento da categorização inicial, mas ainda assim, dentro dessa etapa.

Por fim, reafirma-se a necessidade de expansão das análises de subcorpus para outras pesquisas qualitativas. Assim, será possível maior riqueza colaborativa nas compreensões dos achados. Em particular, com vistas à melhoria desta pesquisa, cabe o trato em detalhes sobre as formações das demais categorias intermediárias apresentadas, além da necessidade de anunciar as categorias finais, porém, declara-se que o esforço aqui foi atender à demanda das análises de subcorpus e apontar novas perspectivas analíticas frente ao uso do método de Análise Textual Discursiva apoiado pelo software IRaMuTeQ.

Agradecimentos

Capex – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pela concessão de bolsa de Pós-Doutoramento a um dos autores (Processo 88882.314867 - PNPd/CAPES).

5. Referências

- Albuquerque, N. F., Lima, V. M. R. & Rosito, B. A. (2016). Clube de Ciências como espaço para desenvolver atributos do domínio intrapessoal. *Areté - Revista Amazônica de Ensino de Ciências*, 9(19), 23-33.
- Bernet, J. T. (2013). La Educación No Formal. In Morales, M. (Org.), *Educación No Formal: Lugar de conocimientos: Selección de textos*. Ministerio de Educación y Cultura. Montevideo: Dirección de Educación del Mec Uruguay. p. 27-50.
- Camargo, B. V., Justo, A. M. (2013). IRAMUTEQ: um software gratuito para análise de dados textuais. *Temas em Psicologia*, 21(2), 513-518.
- Camors, J. (2013). Educación no Formal: Política Educativa del MEC 2005-2009. In Morales, M. (Org.), *Educación No Formal: Lugar de conocimientos: Selección de textos*. Ministerio de Educación y Cultura. Montevideo: Dirección de Educación del Mec Uruguay. p. 51-70.
- Costa, A. P., Amado, J. (2018). *Análise de conteúdo suportada por software*. Aveiro/POR: Ludomedia.

De Prá, G., Tomio, D. (2014). Clube de Ciências: condições de produção da pesquisa em educação científica no Brasil. *Alexandria: Revista de Educação em Ciência e Tecnologia*, Florianópolis, 7(1), 179-207.

Gray, D. (2012). *Pesquisa no mundo real*. Porto Alegre: Penso.

Kami, M. T. M., Larocca, L. M., Chaves, M. M. N., Lowen, I. M. V., Souza, V. M. P., Goto, D. Y.N. (2016). Trabalho no consultório na rua: uso do software IRAMUTEQ no apoio à pesquisa qualitativa. *Escola Anna Nery*, 20(3). DOI: 10.5935/1414-8145.20160069.

Lima, V. M. R. (1998). *Clubes de Ciências: contribuições à formação do educando*. Dissertação (Mestrado em Educação) - Faculdade de Educação Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre.

Lima, V. M. R., & Ramos, M. G. (2017). Percepções de interdisciplinaridade de professores de Ciências e Matemática: um exercício de Análise Textual Discursiva. *Revista Lusófona de Educação*, 36, 163-177.

Mancuso, R., Lima, V. M. R., & Bandeira, V. (1996). *Clubes de Ciências: criação, funcionamento, dinamização*. Porto Alegre: SE/CECIRS.

Martins, I. C. S., Lima, V. M. R., Amaral-Rosa, M. P., Moreira, L., Ramos, M. G. (2020). Handcrafted and Software-Assisted Procedures for Discursive Textual Analysis: Analytical Convergences or Divergences? In Costa, A. Reis, L. Moreira, A. (Eds). *Computer Supported Qualitative Research*. WCQR 2019. Advances in Intelligent Systems and Computing, vol. 1068. Springer, Cham. pp. 189-205.

Mayring, P. (2014). *Qualitative Content Analysis: theoretical foundation, basic procedures and software solution*. Klagenfurt, Austria. Recuperado de: <https://goo.gl/pNjubm>.

Moraes, R. Galiazzi, M. C. & Ramos, M. G. (2013). Aprendentes do aprender: um exercício de Análise Textual Discursiva. *Indagatio Didactica*, 5(2), 868-883.

Moraes, R., & Galiazzi, M. C. (2016). *Análise Textual Discursiva*. Ijuí: Ed. Unijuí.

Ramos, M. G., Lima, V. M. R., & Amaral-Rosa, M. P. (2019). IRAMUTEQ Software and Discursive Textual Analysis: Interpretive Possibilities. In Costa, A., Reis, L., Moreira, A. (Eds.). *Computer Supported Qualitative Research*. WCQR 2018. Advances in Intelligent Systems and Computing, vol. 861. Springer, Cham. Pp. 58-72.

Ratinaud, P. (2014). *IRAMUTEQ: Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires – 0.7 alpha 2*. Recuperado de: <http://www.iramuteq.org>.

Rosito, B. A. & Lima, V. M. R. (2020). *Conversas sobre Clubes de Ciências*. Porto Alegre: EDIPUCRS.

Stake, R. E. (2011). *Pesquisa qualitativa: estudando como as coisas funcionam*. Porto Alegre: Penso.

Tomio, D.; Hermann, A. P. (2019). Mapeamento dos clubes de ciências da América Latina e construção do site da rede internacional de clube de ciências. *Ensaio: Pesquisa em Educação em Ciências*, 21(1), 1-23.

Viêra, M., Lima, V. M. R. (2016). *Entrelaçar de histórias: a vida de um professor e o Centro de Ciências do Rio Grande do Sul*. Curitiba: Appris.

Veraszto, E. V., Camargo, E. P., Camargo, J. T. F., Simon, F. O., Miranda, N. A. (2018). Evaluation of concepts regarding the construction of scientific knowledge by the congenitally blind: an approach using the correspondence analysis method. *Ciênc. Educ.*, 24(4), 837-857.