

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

Scientific production versus research lines of the graduate study programs in Sociology and Social Sciences in Brazil

Velcimiro Inácio Maia¹

Resumo: O objetivo foi mapear as subáreas da produção qualificada e das linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais (PPGS), durante dois triênios. Foi constatada a falta de alinhamento entre as temáticas da produção qualificada e das linhas de pesquisa, confrontando-se os agrupamentos de PPGS com linhas de pesquisa semelhantes com os agrupamentos pelas temáticas da produção bibliográfica. A metodologia consistiu na pesquisa qualitativa por meio da análise lexical e de conteúdo dos artigos com *Qualis* A1, A2 e B1 e das descrições das linhas de pesquisa dos PPGS. Para isso, utilizou-se o *software Iramuteq 0.7 alpha 2*. Foram analisados 53 PPGS e 1627 artigos. Nas considerações finais apontamos algumas hipóteses para a questão do não alinhamento da produção com as linhas de pesquisa e direcionamentos para possíveis desdobramentos da pesquisa.

Palavras-chave: Análise lexical; *Iramuteq*; Pós-graduação;

¹ Professor Adjunto, DTECH-Departamento de Tecnologia e Eng. Civil, Computação e Humanidades/UFSJ. Membro do GIARS-Grupo Interdisciplinar de Pesquisa em Análise de Redes Sociais/UFMG. E-mail: maia@ufs.edu.br

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

Produção científica; Linhas de pesquisa.

Abstract: The objective was to map as subareas of the qualified production and research lines of the graduate study programs in Sociology and Social Sciences, during two periods of three years. There was a lack of alignment between the issues of qualified production and research lines.

This was achieved by comparing the clusters of similar research lines with the clusters of bibliographic production themes. Qualitative and quantitative research was used through the lexical and content analysis of papers with Qualis A1, A2 and B1 and descriptions of the research lines of the graduate programs. For this, the software Iramuteq 0.7 alpha 2 was used. We analyzed 53 programs and 1627 papers. The conclusion showed some hypotheses for the non-alignment of scientific production with the lines of research, as well as directions for possible research developments.

Keywords: Lexical analysis; Iramuteq; graduate programs; Scientific production; Research lines.

1 Introdução

O aumento da produção científica e de sua qualidade tem sido uma das grandes metas da pós-graduação no Brasil e no mundo. Vários meios de divulgação surgiram e ampliaram a cobertura dos periódicos, SciELO e *Web of Science* são exemplos de indexadores de grande abrangência temática. Em junho de 2016 existiam 56 bases de indexação na área de Ciências Humanas, subárea Sociologia, cadastradas no portal de periódicos Comissão de Aperfeiçoamento de Pessoal de Nível Superior (Capes). A publicação de resultados tornou-se ato inseparável da atividade de pesquisa e é por meio das publicações que um campo científico adquire maturidade e se torna visível, tal maturidade pode ser estudada a partir da

análise da produção científica, ao mesmo tempo em que possibilita o mapeamento de uma disciplina ao conhecer as temáticas pesquisadas (Machado, 2013).

A Capes é responsável pelo reconhecimento e avaliação de cursos de pós-graduação *stricto-sensu* em âmbito nacional. A avaliação trienal dos cursos é desenvolvida por comitês de área, formados por consultores acadêmicos escolhidos dentre profissionais de comprovada competência em pesquisa e ensino de pós-graduação. Os comitês também são responsáveis pelo estabelecimento de critérios e diretrizes para a avaliação, que são divulgados em documentos das respectivas áreas. Compete aos comitês a qualificação de periódicos, anais, jornais, revistas e livros (*Qualis*). O *Qualis* é um levantamento realizado com o intuito de mensurar a qualidade da produção científica dos programas de pós-graduação. Na verdade, avalia a qualidade dos periódicos em que eles são publicados. A classificação de periódicos é atualizada anualmente. Esses veículos são enquadrados em estratos indicativos de qualidade – A1(mais elevado); A2; B1; B2; B3; B4; B5 e C.

Por outro lado, podemos conjecturar que os Programas de Pós-Graduação em Sociologia (PPGS) deveriam estabelecer suas linhas de pesquisa de acordo com as possibilidades de maximização de seus recursos humanos e materiais. Diante destas considerações, traçamos dois objetivos: (1) mapear as subáreas temáticas refletidas pela produção mais qualificada, assim como as subáreas segundo as propostas dos PPGS em suas linhas de pesquisa e (2) verificar se os agrupamentos de programas segundo a similitude de temas da produção qualificada coincidem com os agrupamentos segundo a similitude das linhas de pesquisa. Assim, este estudo desvelou os subcampos científicos que traduzem as temáticas trabalhadas pelos diferentes pesquisadores em seus PPGS.

2 Procedimentos metodológicos

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

Esta pesquisa baseou-se na análise quali-quantitativa de material textual. Optamos por utilizar a análise lexical além da análise de conteúdo, dado o grande volume de material a ser analisado e os objetivos da pesquisa. De acordo com Bardin (1977), a análise de conteúdo é definida como:

Um conjunto de técnicas de análise das comunicações visando obter, por procedimentos, sistemáticos e objetivos de descrição do conteúdo das mensagens, indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) destas mensagens. (p. 44).

A análise de conteúdo surgiu como uma tentativa de sistematização de dados qualitativos. Conforme apontam Lebart e Salem (1994)² apud Justo e Camargo (2014), a análise de conteúdo opera em dois tempos. O analista começa por definir um conjunto de classes de equivalência – categorias que podem ser definidas a priori ou a posteriori, as quais terão suas ocorrências identificadas ao longo do texto a ser analisado. Em um segundo momento, é feita a contagem de cada uma dessas categorias e a construção de tabelas que sistematizam a análise. Portanto, a experiência do pesquisador é crucial para a correta classificação e interpretação dos dados. Justamente por isso, as pesquisas qualitativas dadas a cabo desta forma costumam ser muito criticadas, vista a exposição à subjetividade e interpretação do pesquisador.

O volume de material analisado pode tornar o método da análise de conteúdo impraticável. Imagine um pesquisador tendo que analisar um conjunto de mais de mil entrevistas, ou

² LEBART, L.; SALEM, A. *Statistique textuelle*. Paris: Dunod, 1994.

algum outro tipo de material textual mais extenso, como artigos científicos, por exemplo.

Considera-se que existem alguns tipos de dados que dificultam a análise de conteúdo, por se tratarem de bancos de dados demasiado extensos, os quais as equipas de investigadores teriam dificuldade em analisar de modo coeso e fidedigno, ou mesmo que este trabalho despenderia de um tempo que é incompatível com os atuais prazos relativos aos estudos de mestrado, doutorado ou editais de pesquisa (Justo; Camargo, 2014, p.4)

Limitar a análise a pequenos casos ou trechos textuais sobre os quais o pesquisador julga suficientes para uma representação do todo não faz com que as críticas diminuam, justamente pela dificuldade de se estabelecer critérios objetivos para esta escolha. A análise lexical de dados textuais surge então como uma alternativa aos estudos qualitativos.

Este modelo se mostra mais adequado para a análise da clusterização da produção científica por áreas temáticas e pelos programas de pós-graduação, ou seja, encontrar os subgrupos de programas que tratam sobre os mesmos 'assuntos'. Lahlou (1994)³ apud Justo e Camargo (2014), considera que as análises lexicais viabilizam a superação da dicotomia clássica entre quantitativo e qualitativo na análise de dados, na medida em que a partir delas se torna possível quantificar e empregar cálculos estatísticos sobre variáveis essencialmente qualitativas – os textos. Os materiais textuais são os dados qualitativos mais comumente utilizados. Estes dados podem ser provenientes de entrevistas, questionários, grupos de discussões, atas, redações

³ LAHLOU, S. *L'analyse lexicale. Variances*, v.3, p. 13-24. 1994.

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

etc. Enfim, há uma grande variedade de possibilidades para se efetuar a coleta de dados qualitativos na forma textual.

A análise lexical tornou-se uma nova fronteira entre a pesquisa puramente qualitativa e as essencialmente quantitativas. O método procura manter as vantagens das duas abordagens, ao mesmo tempo em que minimiza os pontos fracos de cada uma. A pesquisa quantitativa, com dados coletados por meio de questionários fechados, por exemplo, sempre corre o risco de não ser abrangente o suficiente para abarcar todas as realidades sobre um fato ou objeto estudado (o pesquisador sempre terá que limitar de alguma forma as alternativas de resposta). Por outro lado, a coleta de dados por meio de questões abertas ou entrevistas poderia gerar respostas tão abrangentes a ponto de dificultar a correta interpretação.

Na análise lexical o material a ser analisado deve ser preparado em um *corpus* textual que inclui os textos a analisar e sobre o qual o pesquisador poderá atribuir uma ou mais variáveis.

Torna-se possível, a partir da análise lexical, descrever um material produzido por determinado produtor, seja individual ou coletivamente (um indivíduo ou um grupo), como também pode ser utilizado este tipo de análise com a finalidade comparativa, relacional, comparando produções textuais diferentes em função de variáveis específicas que descrevem quem produziu o texto. O objetivo é, a partir dessas técnicas, combinar vantagens dos enfoques quantitativos e qualitativos sem acumular as suas limitações. (Justo; Camargo, 2014, p.4)

As variáveis atribuídas aos textos podem ser cruzadas de forma a permitir a verificação de hipóteses. O pesquisador pode verificar, por exemplo, se homens ou mulheres têm opiniões

diferentes sobre determinado assunto, desde que tenha atribuído tais variáveis por ocasião da codificação do *corpus* textual. No caso desta pesquisa, o interesse foi verificar os agrupamentos de PPGS que se formam segundo as temáticas de pesquisa. A análise lexical dos *corpora* textuais preparados a partir dos grupos de palavras-chave dos artigos mais qualificados produzidos pelos PPGS e de suas linhas de pesquisa permitiu ao mesmo tempo desvelar as subáreas de pesquisa e os PPGS que trabalham em torno de temas semelhantes.

Lançando o olhar sobre a produção bibliográfica publicada nos periódicos mais qualificados, buscamos mostrar como o campo da sociologia se divide em 'subáreas temáticas'. Uma opção trivial seria a utilização de alguma 'taxonomia' das sociologias, a fim de enquadrar as publicações dos pesquisadores segundo as 'subáreas', como a tabela de áreas do conhecimento da Capes, por exemplo. Talvez esta alternativa fosse produtiva para uma classificação nas chamadas *hard sciences*, no entanto, entendemos que tal forma de classificação não seria tão produtiva quanto a que escolhemos para o caso do campo da sociologia. Destarte, nossa opção foi fazer emergir da própria produção bibliográfica tais subáreas temáticas, como subcampos do conhecimento científico gerado pelos PPGS no período de dois triênios. Em vez de lançar mão de uma tabela pré-definida e buscar enquadrar cada artigo publicado em determinada subárea, optamos por uma abordagem da análise lexical. Procuramos, por meio do processamento das palavras-chave dos artigos publicados, definir os agrupamentos de termos semelhantes, classificando-os posteriormente em subáreas temáticas.

O primeiro passo foi levantar todos os artigos dos triênios de 2007 a 2009 e 2010 a 2012 publicados em periódicos com *Qualis* A1, A2 e B1. Este corte foi definido com a finalidade de filtrar as publicações mais qualificadas, tornando a coleta mais eficiente, visto a grande quantidade de artigos na área da Sociologia e

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

Ciências Sociais. Em segundo lugar, analisamos as linhas de pesquisa propostas pelos programas durante estes triênios. Em ambos os casos, foi usado um método de pesquisa baseado tanto na análise quantitativa de dados textuais, quanto na análise qualitativa de conteúdo. Procuramos desenvolver uma análise sistemática e objetiva dos dados textuais obtidos da produção bibliográfica qualificada durante os dois triênios, assim como das linhas de pesquisa. “Uma análise quantitativa de dados textuais não deixa de considerar a qualidade do fenômeno estudado, e ainda oferece critérios provenientes do próprio material para a consideração do mesmo como indicador de um fenômeno de interesse científico” (Camargo, 2005, p. 511).

Por sorte, o desenvolvimento da lexicografia, da informática e da computação nas últimas décadas possibilitou avanços nas análises quantitativas de dados textuais. Um software adequado para tais estudos, o *Alceste (Analyse Lexicale par Contexte d'un Ensemble de Segment de Texte)*, foi desenvolvido pela equipe de John Paul Benzécri no *Centre National de la Recherche Scientifique-CNRS* na França, a partir da obra de Max Reinert e é muito utilizado desde a década de 1990 em pesquisas nas áreas de Sociologia e Psicologia, nas quais se faz necessária a análise de material textual.

No entanto, nesta pesquisa utilizamos o software *Iramuteq 0.7 alpha 2 (Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires)* que utiliza o mesmo algorítmico do *Alceste* e usa a análise de Classificação Hierárquica Descendente (CHD). O *software* permite a análise lexicográfica do material textual, gerando contextos (classes lexicais) que são caracterizados pelo seu vocabulário e pelos segmentos de texto que compartilham este vocabulário (Reinert, 1998). Seu método tem como base a organização da distribuição do vocabulário de forma compreensível e visualmente clara.

O primeiro passo na utilização do *software*, e também o mais trabalhoso, é a preparação de um *corpus* textual a ser

analisado. No nosso caso, preparamos dois *corpora* textuais para cada triênio: um *corpus* com os segmentos de palavras-chave dos artigos qualificados e outro com as linhas de pesquisa, totalizando quatro *corpora* analisados. Com o intuito de deixar claro para o leitor, neste ponto daremos algumas definições básicas para a utilização do *Iramuteq*:

Corpus: banco de dados de texto que será analisado.

UCI – Unidade de Contexto Inicial: Unidade a partir da qual é feita a fragmentação inicial. Pode ser um capítulo de livro, uma matéria jornalística, uma resposta de entrevista.

UCE – Unidade de Contexto Elementar: Definida segundo critérios de tamanho do texto (número de palavras analisadas) e pontuação. “É a partir do pertencimento das palavras de um texto a uma UCE, que o programa Alceste vai estabelecer as matrizes a partir das quais será efetuado o trabalho de classificação” (REINERT, 1998, p. 17).

UC – Unidade de Contexto: Agrupamento de UCE sucessivas dentro de uma mesma UCI, até que o número de palavras diferentes analisadas (contidas nessa unidade de contexto) seja superior ao limiar fixado na análise. (...) Esse limiar é fixado proporcionalmente ao número de palavras analisadas. As UC devem ter uma homogeneidade semântica e devem ser de tamanho comparável. Ou seja, são unidades estatísticas de pesos idênticos. Elas estão na base de todas as estatísticas efetuadas (Reinert, 1998, p. 18).

Classe: pode ser definida como um agrupamento constituído por várias UCE de vocabulário homogêneo.

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

“Lematização”: operação de substituição de certas palavras por uma forma reduzida, que permite ao programa considerar equivalentes palavras com radical comum, que significam aproximadamente a mesma coisa, mas que diferem quanto ao gênero, ao número ou ao fato de serem substantivos, adjetivos ou advérbios. Por exemplo: AMOR, AMORES, AMOROSOS, são substituídas no texto pela forma AMOR+.

AFC – Análise Fatorial de Correspondência: Cruzamento entre o vocabulário (considerando a frequência de incidência de palavras) e as classes, gerando uma representação gráfica em plano cartesiano, na qual são vistas as oposições entre classes ou formas. (Reinert, 1998 apud Nascimento; Menandro, 2006, p.75).

Dentre as análises que o programa realiza, a análise lexical clássica é feita de forma que o programa identifica e reformata as unidades de texto, transformando as Unidades de Contexto Inicial (UCI) em Unidades de Contexto Elementares (UCE), identificando a quantidade de palavras e as frequências, criando um dicionário de palavras reduzidas, além de distinguir formas ativas e suplementares. Outra análise, a de especificidade, associa os textos do banco de dados a variáveis descritoras dos seus produtores, como uma análise de contrastes, que é feita dividindo o *corpus* em função de uma variável escolhida pelo pesquisador. Já o método da Classificação Hierárquica Descendente (CHD), produzido por Reinert (1990), classifica os segmentos de texto em função dos seus respectivos vocabulários, repartidos em conjuntos, determinados pela frequência das formas reduzidas, chamados “*lemmes*”. Assim, obtêm-se classes de UCE que podem, ao mesmo tempo, apresentar vocabulário semelhante entre si e diferentes das UCE

de outras classes. Os resultados também podem ser apresentados visualmente em eixos cartesianos, essa análise permite que se resgate do texto os segmentos associados de cada classe.

Após estas análises quantitativas (lexicais) foram realizadas as análises qualitativas (de conteúdo) do material textual, possibilitando o mapeamento da produção qualificada e a sua confrontação com as propostas dos PPGS em suas linhas de pesquisa.

3 Mapeamento das subáreas da produção mais qualificada

A coleta de dados primários para esta pesquisa se deu inicialmente pela identificação dos artigos com *Qualis* A1, A2 e B1 a partir das listas disponibilizadas pela Capes no item 'Publicação Bibliográfica' dos 'Cadernos de Indicadores' em seu *site* da Internet, referentes aos triênios 2007-2009 e 2010-2012 dos PPGS listados no Quadro 1. A área continha 42 programas no primeiro triênio e 52 no segundo.

Quadro 1 – Lista de instituições e programas de pós-graduação por triênios

N.	Instituição	PPGS	Nota Capes Triênios		N.	Instituição	PPGS	Nota Capes Triênios	
			1º	2º				1º	2º
1	FUFPI	Sociologia	-	3	27	UFPA	Defesa Social e Mediação de Conflitos	-	3
2	FUFSE	Ciências Sociais	4	4	28	UFPB/J.P.	Sociologia	3	4

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

3	PUC-RIO	Ciências Sociais	4	4	29	UFPE	Sociologia	5	5
4	PUC/MG	Ciências Sociais	4	5	30	UFPEL	Sociologia	3	3
5	PUC/RS	Ciências Sociais	4	4	31	UFPR	Sociologia	5	4
6	PUC/SP	Ciências Sociais	5	4	32	UFRB	Ciê. Soc. Cultura, Desigualdades e Desenvolvimento	-	3
7	UCAM	Sociologia	7	1	33	UFRGS	Sociologia	6	7
8	UECE	Planejamento e Políticas Públicas	4	5	34	UFRJ	Sociologia e Antropologia	7	7
9	UECE	Políticas Públicas e Sociedade	3	4	35	UFRN	Ciências Sociais	4	4
10	UEL	Ciências Sociais	3	4	36	UFRRJ	Ciências Sociais	-	3
11	UEM	Ciências Sociais	3	3	37	UFRRJ	Ciências Sociais em Desenv., Agricultura e Sociedade	5	5
12	UENF	Sociologia Política	4	3	38	UFSC	Sociologia Política	5	5
13	UERJ	Sociologia	-	5	39	UFSCAR	Sociologia	5	6
14	UERJ	Ciências Sociais	5	5	40	UFSM	Ciências Sociais	3	3
15	UFAL	Sociologia	3	3	41	UFU	Ciências Sociais	-	3
16	UFAM	Sociologia	3	3	42	UNB	Sociologia	5	6
17	UFBA	Ciências Sociais	5	5	43	UNESP/ARAR	Ciências Sociais	4	4
18	UFC	Sociologia	5	5	44	UNESP/MAR	Ciências Sociais	4	4
19	UFCE	Ciências Sociais	4	3	45	UNICAMP	Ciências Sociais	5	5
20	UFES	Ciências Sociais	3	3	46	UNICAMP	Sociologia	6	6
21	UFF	Sociologia	-	3	47	UNIFESP	Ciências Sociais	-	3
22	UFG	Sociologia	4	4	48	UNIOESTE	Ciências Sociais	-	3
23	UFJF	Ciências Sociais	4	4	49	UNISINOS	Ciências Sociais	4	5
24	UFMA	Ciências Sociais	4	4	50	USP	Sociologia	7	7
25	UFMG	Sociologia	4	5	51	UVV	Ciências Sociais	-	3
26	UFPA	Ciências Sociais	4	4	52	UVV	Sociologia Política	-	3
-	-	-	-	-	53	UFMG	Sociologia Política*	3	-

Nota: * Programa de pós-graduação extinto em 2010

Fonte: Elaborado pelo autor a partir dos documentos de Área Capes, 2013

Posteriormente, buscou-se levantar cada artigo listado e preparar um *'corpus'* textual para cada triênio contendo os segmentos de *'palavras-chave'* de cada artigo por PPGS. Para o primeiro triênio foram listados 764 artigos com *Qualis* A1, A2 e B1 pela Capes. Entretanto, foi possível coletar um montante de

547 artigos, o que equivale a 71,60% do total declarado pelos programas na coleta de dados da Capes. Para o triênio 2010-2012 o total de artigos publicados foi de 1.363, sendo possível coletar 1.080 artigos (74,69%). As diferenças de 217 documentos, para o primeiro triênio, e de 345 para o segundo, foram devidas ao fato de que muitos documentos na verdade não se apresentavam como 'artigos científicos', mas às vezes como resenhas, apresentações ou mesmo estavam sob a forma de artigos, mas não continham palavras-chave de forma destacada no texto original. Houve ainda casos em que duas publicações de conteúdo idêntico se encontravam listadas repetidamente dentro do mesmo programa em anos diferentes, mas dentro do mesmo triênio, estas duplicidades foram excluídas. A Tabela 1 sintetiza as quantidades de artigos coletados.

Tabela 1 – Total de artigos *Qualis* A1, A2 e B1 coletados

Triênios	Total de artigos listados pela Capes	Artigos coletados	Percentual
2007-2009	764	547	71,60%
2010-2012	1.363	1.018	74,69%

Fonte: Elaborado pelo autor

Como o material analisado foi constituído pelos grupos de palavras-chave das publicações, os materiais que não apresentavam este item em seus textos foram desconsiderados, lembrando que os materiais escritos em idioma diferente do Português foram traduzidos, visto que foi definida no software a utilização do dicionário em Português.

Concluído o banco de dados com os artigos qualificados dos programas de pós-graduação, preparamos o *corpus* textual, considerando os grupos de palavras-chave de cada publicação. O *Iramuteq* exige uma preparação do *corpus* de forma que exista entre cada texto uma linha de codificação por asteriscos, estas codificações indicarão ao software as variáveis independentes

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

que serão cruzadas com as variáveis dependentes (palavras-chave neste caso).

Cada segmento de texto foi codificado da seguinte forma:

```
**** *art_0001 *progr_2
Africanos no Brasil, Identidades,
Solidariedades, Mobilidades, Táticas,
Sergipe.
**** *art_0002 *progr_2
Revolução, Social Democracia, Movimento
dos Operários, Europa.
[...]
```

A primeira codificação (*art_0001) indica a numeração atribuída a cada artigo coletado, a segunda codificação (*progr_2) indica a qual PPGS a publicação pertence.

O programa *Iramuteq* viabiliza vários tipos de análises com dados textuais, dentre as mais simples, como a lexicografia básica, que consiste no cálculo de frequência de palavras, até análises multivariadas, como a Classificação Hierárquica Descendente (CHD), e a análise de similitude. Neste estudo, concentramos a análise da lexicografia e da CHD. Em cada análise realizada, o software executa uma lematização.

A lematização é realizada a partir do dicionário do *Iramuteq*, sem desambiguação, os verbos são reduzidos ao infinitivo, os nomes ao singular e os adjetivos ao masculino singular. Para cada nova análise, o programa questiona se a lematização será mantida. Além da classificação dos lemas, a análise lexicográfica realiza também a contagem do número de formas ativas (classes gramaticais), número de formas suplementares, número de classes, entre outras. Os segmentos de texto são classificados através do método de CHD, em função dos seus respectivos vocabulários, e o agrupamento deles é dividido em função da frequência das formas reduzidas. Ao aplicar o método de CHD, a partir de matrizes, cruzando

segmentos de textos e palavras (em repetidos testes do tipo X^2), uma classificação estável e definitiva é obtida (Reinert, 1990).

O software organiza a análise dos dados em um dendrograma da CHD, por meio das análises em matrizes, ilustrando as relações entre as classes. A descrição de cada uma das classes é fornecida através dos cálculos e resultados que o programa executa, principalmente pelo seu vocabulário característico (léxico) e pelas variáveis definidas no *corpus*.

3.1 Análise lexical da publicação científica

A análise lexicográfica básica do *corpus* textual, preparado com os segmentos de palavras-chave dos artigos do primeiro triênio, resultou nos seguintes dados:

Número de textos: 547 (segmentos, grupos de palavras-chave)

Numero de ocorrências: 4514 (total de palavras)

Número de formas: 1526 (formas distintas)

Número de lemas (*lemmes*): 1320 (formas reduzidas)

Número de formas ativas: 1289 (classes gramaticais escolhidas: nomes e adjetivos)

Número de formas suplementares: 26 (classes gramaticais escolhidas: preposições, artigos def., advérbios)

Número de formas ativas com frequência ≥ 3 : 331

Média de formas por segmento: 8,25

Número de classes: 3

427 textos classificados de 547 (78,06 %)

Os algoritmos empregam o método de análise CHD e gera classes de segmentos de texto que, ao mesmo tempo,

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

apresentam vocabulário semelhante entre si, e vocabulário diferente dos segmentos de texto das outras classes.

O *método da Classificação Hierárquica Descendente* (CHD) proposto por Reinert (1990) e utilizado pelo *software Alceste* classifica os segmentos de texto em função dos seus respectivos vocabulários, e o conjunto deles é repartido com base na frequência das formas reduzidas (palavras já lematizadas). Esta análise visa obter classes de UCE que, ao mesmo tempo, apresentam vocabulário semelhante entre si, e vocabulário diferente das UCE das outras classes. O *IRAMUTEQ* também fornece outra forma de apresentação dos resultados, por meio de uma análise fatorial de correspondência feita a partir da CHD (Análise Pós-Fatorial) que representa num plano cartesiano as diferentes palavras e variáveis associadas a cada uma das classes da CHD. A interface possibilita que se recuperem, no *corpus* original, os segmentos de texto associados a cada classe, momento em que se obtém o contexto das palavras estatisticamente significativas, possibilitando uma análise mais qualitativa dos dados. (Camargo; Justo, 2013, p. 516).

Para o triênio 2007-2009, dos 547 segmentos de texto analisados, que equivalem aos conjuntos de palavras-chave dos artigos coletados, 427 foram classificados. O resultado desta análise CHD apresentou três classes lexicais, como pode ser visualizado na Figura 1.

Para que o leitor compreenda os resultados que serão apresentados e com a finalidade de transparência quanto ao

método utilizado, exibimos os detalhes definidos no *software*. Os parâmetros utilizados no *Iramuteq* em 'Especificidades e AFC' foram: lematizado e indexado pelo dicionário em Português⁴; formas utilizadas: ativas e suplementares; variável escolhida para o teste: programas de pós-graduação; distribuição: Qui-quadrado; formas com frequência mínima de 3 ocorrências. Na Classificação Hierárquica Descendente (CHD), pelo método Reinert, utilizamos os seguintes parâmetros: lematizado e indexado com o dicionário; número máximo de formas analisadas = 3.000.

⁴ O dicionário conta com um léxico de 142.598 vocábulos e foi elaborado pelo Laboratório de Psicologia Social da Comunicação e Cognição da Universidade Federal de Santa Catarina (LACCOS/UFSC) em parceria com o CIERS-ed/FCC e com o grupo de pesquisa Valores, Educação e Formação de Professores da UNESP. (Camargo; Justo, 2013, p.516).

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

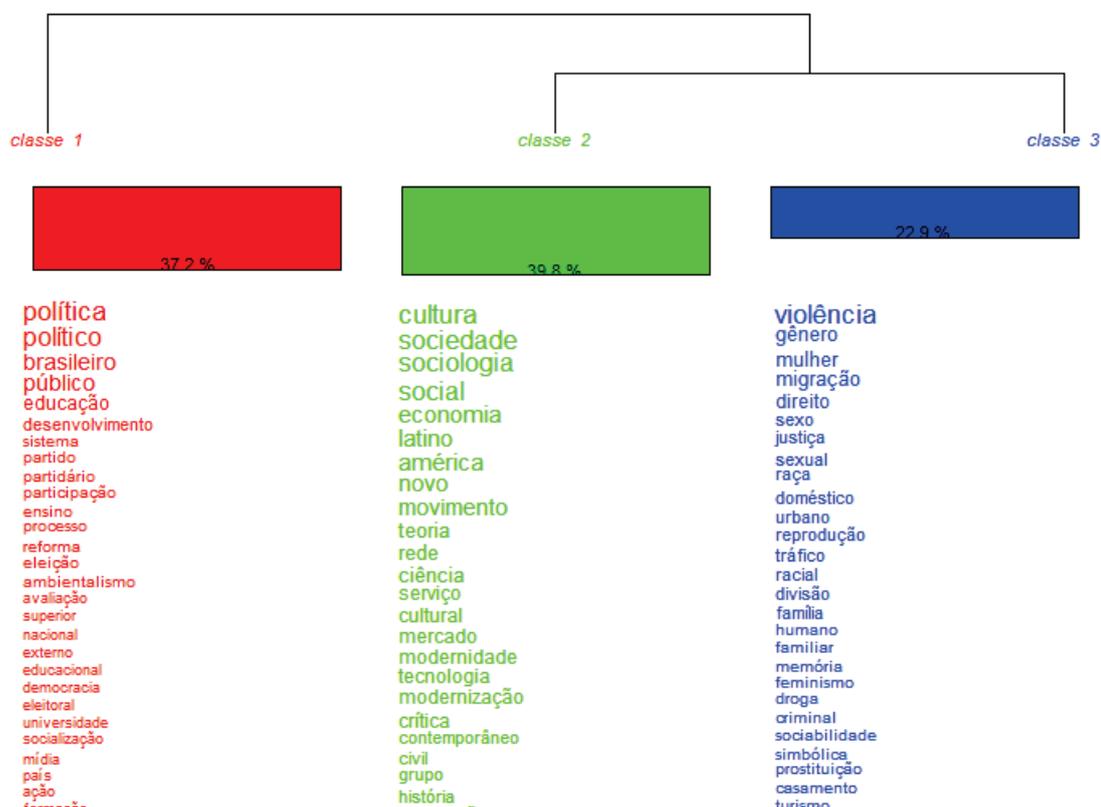


Figura 1 – Dendrograma das classes triênio 2007-2009 segundo produção qualificada

Fonte: corpus processado em Iramuteq

A classe 1 agrupou 159 dos 427 segmentos de texto classificados, o que equivale a 37,24% do total de artigos científicos considerados na classificação. A classe 2 agrupou 170 segmentos, ou seja, 39,81% do total de artigos classificados. A classe 3 agrupou 98 segmentos, ou seja, 22,95% do total de artigos classificados.

Para o segundo triênio, a análise dos segmentos de palavras-chave resultou nos seguintes dados:

Número de textos: 1018 (segmentos, grupos de palavras-chave)

Numero de ocorrências: 8072 (total de palavras)
Número de formas: 2264 (formas distintas)
Número de lemas (*lemmes*): 1936 (formas reduzidas)
Número de formas ativas: 1885 (classes gramaticais escolhidas: nomes e adjetivos)
Número de formas suplementares: 41 (classes gramaticais escolhidas: preposições, artigos def., advérbios)
Número de formas ativas com frequência ≥ 3 : 535
Média de formas por segmento: 7,93
Número de classes: 4
848 textos classificados de 1018 (83,30 %)

Para o triênio 2010-2012, dos 1.018 segmentos de texto analisados, que equivalem aos conjuntos de palavras-chave dos artigos coletados, 848 foram classificados. O resultado da análise CHD apresentou quatro classes lexicais, como pode ser visualizado na Figura 2.

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

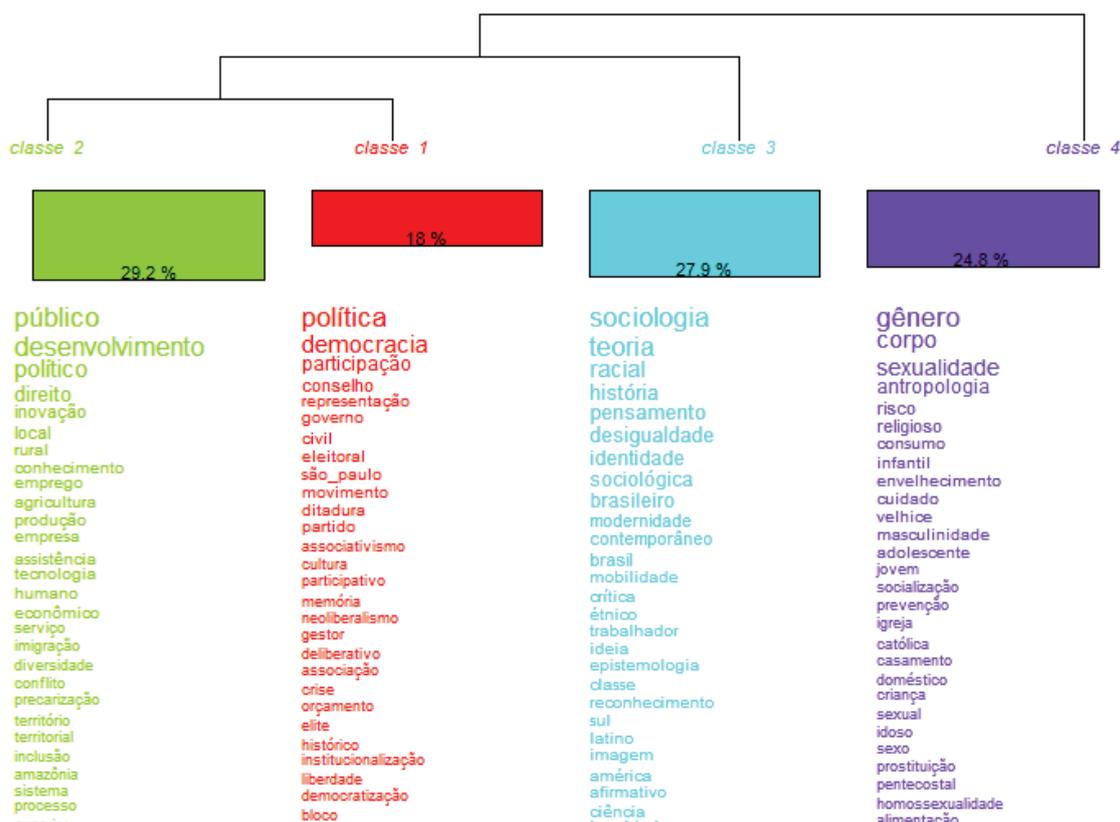


Figura 2 – Dendrograma das classes triênio 2010-2012 segundo produção qualificada

Fonte: corpus processado em Iramuteq

A classe 1 agrupou 153 dos 848 segmentos de texto classificados, o que equivale a 18,04% do total de artigos científicos considerados na classificação. A classe 2 agrupou 248 segmentos, ou seja, 29,25% do total de artigos classificados. A classe 3 agrupou 237 segmentos, ou seja, 27,95% do total de artigos classificados. Por fim, a classe 4 agrupou 210 segmentos, ou seja, 24,76% do total.

Outra forma de apresentação dos resultados é fornecida pelo programa por meio da análise fatorial de correspondência feita a partir da CHD. Os segmentos de texto mais característicos

de cada classe são calculados e fornecidos com base nas classes escolhidas, possibilitando a contextualização do vocabulário típico de cada classe (Fig. 3 e 4). Além de fornecer um *corpus* com os segmentos separados por cores diferentes de acordo com as classes, o software também exibe em cores os lemas mais frequentes por classe no plano cartesiano.

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

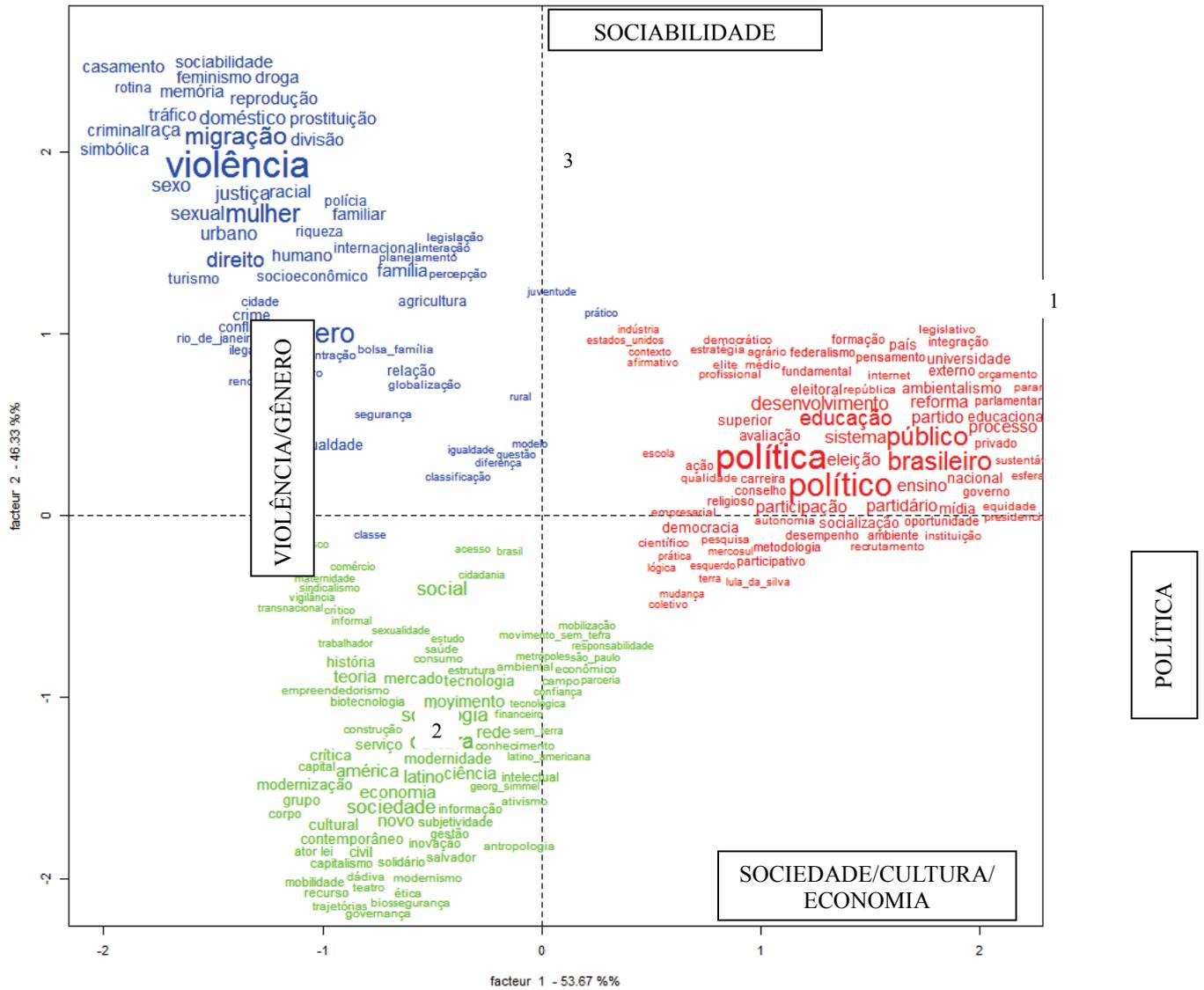


Figura 3 – Três classes da Análise Fatorial de Correspondência triênio 2007-2009 segundo produção qualificada
 Fonte: corpus processado em Iramuteq

A Análise Fatorial de Correspondência (AFC) realiza os cruzamentos entre as formas e as variáveis, lembrando que a variável considerada no *corpus* textual foi o programa de pós-graduação de afiliação dos autores dos artigos. Com isso, foi possível identificar os programas de pós-graduação que publicam segundo áreas temáticas semelhantes.

Análise Fatorial de Correspondência nos dá a seguinte intuição: quanto mais distantes os elementos se encontram dispostos no plano menos eles representam as mesmas coisas. A distância entre as classes no eixo das abscissas (*facteur 1*) responderá por uma porção da distinção entre as classes quanto à determinadas temáticas e o eixo das ordenadas (*facteur 2*) responderá por outra porção da distinção quanto a outras temáticas.

Na Figura 4 o eixo cartesiano correspondente ao '*facteur1*' responde por 53,67% do distanciamento entre as classes, isto é, o quanto um tema diz coisas diferentes e independentes do outro, enquanto o '*facteur 2*' responde pelos 46,33% restantes. Uma interpretação qualitativa dos dados permite ao pesquisador nomear os eixos em extremos semânticos (nesta etapa consultamos as palavras-chave e os resumos dos artigos em cada classe). Para o primeiro triênio atribuímos ao eixo '*facteur 1*' dois polos: Temas relacionados à Violência/Gênero (esquerda) – Temas relacionados à Política (direita); ao eixo '*facteur2*' os polos: Temas relacionados a sociabilidade (acima) – Temas relacionados à Sociedade e Cultura (abaixo). Podemos dizer que o eixo horizontal corresponde a um *continuum* referente à dominação – em um extremo teríamos a dominação em sua esfera privada e no outro à dominação em sua esfera pública. No eixo vertical teríamos um *continuum* referente ao aspecto estrutural, em um extremo os aspectos mais ligados à esfera econômica e no outro, aspectos ligados mais à esfera doméstica, particular.

Observando a Figura 3, referente à Análise Fatorial de

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

Correspondência e o Dendrograma do triênio 2007-2009, pode-se constatar que os dois blocos formados pelas classes 2 e 3 possuem maior proximidade entre si. Já o bloco formado pela classe 1, encontra-se mais distanciado dos outros, mostrando ser uma área de pesquisa mais independente.

A classe 1 tem como destaque o tema político, e associado a ele estão as palavras: política, democracia, participação etc. Pelo dendrograma, essa classe encontra-se menos vinculada às demais. A classe 2 é constituída pelas UCE relacionadas ao tema cultura e sociedade, tais como: cultura, sociedade, sociologia, social, economia etc. Já a classe 3 é constituída pelas palavras relacionadas ao tema violência, tais como: violência, gênero, mulher, migração etc. Pelo dendrograma (Fig.1) podemos observar a proximidade dessas duas classes.

Para o triênio 2010-2012 (Fig. 4), atribuímos ao eixo '*facteur 1*' dois polos: Política/Democracia (esquerda) – Gênero/Religião (direita); ao eixo '*facteur 2*' os polos: Desenvolvimento social/econômico/ambiental (acima) – Teoria social (abaixo). Os *facteurs* 1 e 2, respondem por 39,7% e 32,34% do distanciamento horizontal e vertical, respectivamente. Um terceiro fator, não representado no gráfico, responde por 27,96% do distanciamento. No eixo horizontal correspondente ao *continuum* da dominação temos agora uma configuração invertida, em um extremo a esfera pública (aspectos referentes à política e democracia) e no outro a dominação em sua esfera privada (aspectos ligados à questão de gênero, sexualidade etc., mas incluindo agora estudos sobre Religião). No eixo vertical teríamos um *continuum* referente ao aspecto estrutural, em um extremo os aspectos ligados à esfera do desenvolvimento socioeconômico e ambiental e no outro os aspectos ligados aos estudos sociológicos mais próximos ao campo da teoria social.

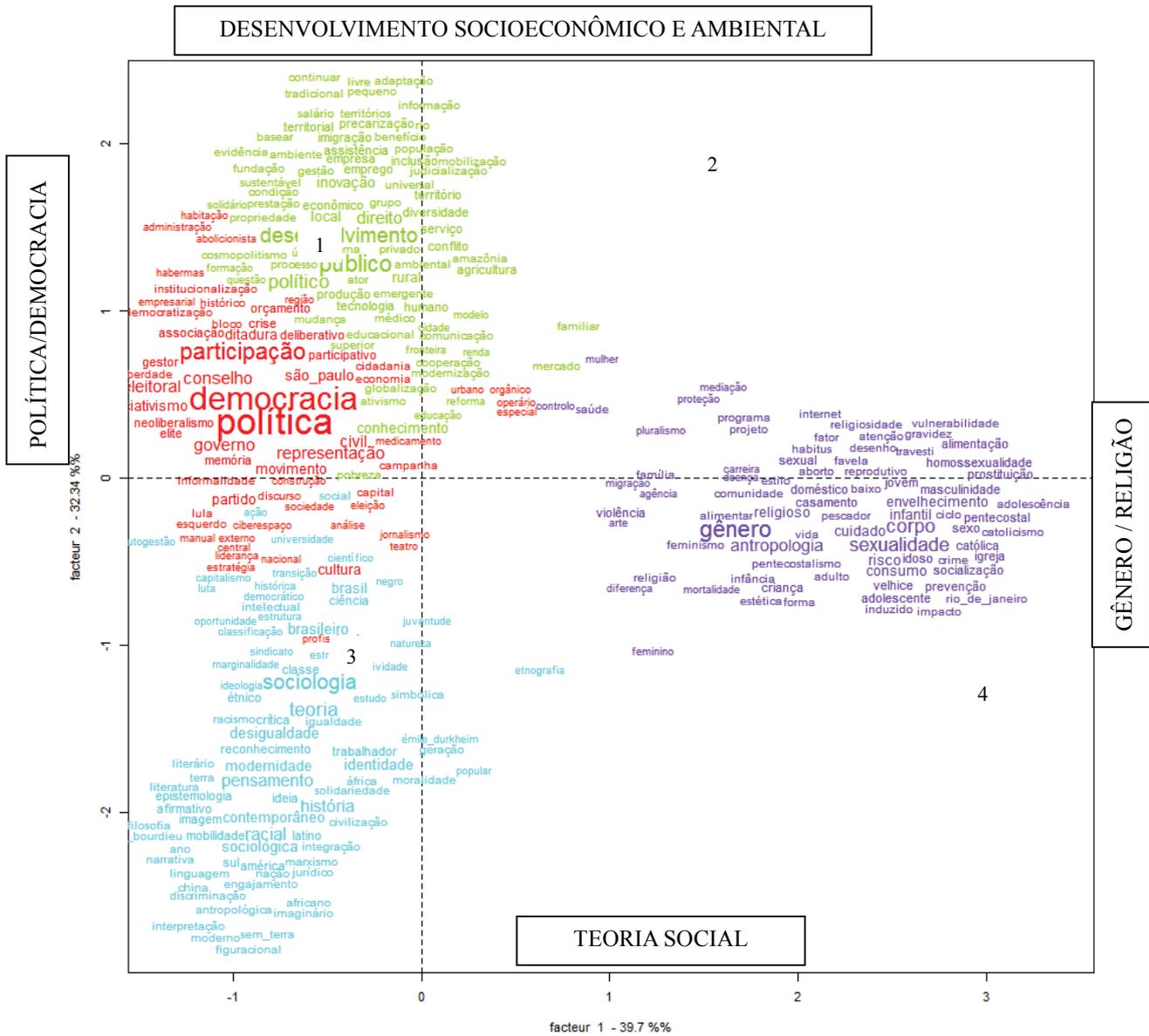


Figura 4 – Quatro classes da Análise Fatorial de Correspondência triênio 2010-2012 segundo produção qualificada

Fonte: corpus processado em Iramuteq

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

Pode-se perceber como as classes 1 e 2 são bastante interseccionadas, isto é, não possuem uma fronteira clara e precisa, seus temas se sobrepõem de certa forma até conseguirem se distanciar. A classe 1, com menor representação na produção qualificada, se situa entre as fronteiras das classes 2 e 3. Na sequência, a classe 3 se torna mais independente das classes 1 e 2. A classe 4, por sua vez, se encontra mais distante, mais independente das outras, mostrando ser uma área de pesquisa mais independente das outras três áreas.

A classe 1 é constituída pelas palavras relacionadas à temática política e de participação pública, tais como: política, democracia, participação e conselho. Já a classe 2 é constituída por palavras relacionadas às temáticas ligadas às questões relacionadas ao desenvolvimento social e econômico, tais como: público, desenvolvimento, político, direito, inovação, local e rural. A classe 3 tem constitui a temática ligada ao campo de cunho mais teórico da sociologia, as palavras mais relacionadas são: sociologia, teoria, racial, história, pensamento. A classe 4 é constituída pelas palavras mais relacionadas com a temática gênero, sexualidade e religião, como: gênero, corpo, sexualidade, antropologia e religioso. Através do dendrograma podemos observar a proximidade entre as classes 1 e 2 e, posteriormente, 1 e 2 com a classe 3 e a classe 4 mais afastada.

De um triênio ao outro houve aumento da diversidade de produção científica da área. Os dados apresentados acima mostraram como o campo se tornou mais complexo. As classes 1 dos dois triênios continuaram tratando sobre as temáticas relacionadas à política, democracia e participação. As classes 3 e 4 do primeiro e segundo triênios, respectivamente, trataram sobre gênero e sexualidade, no entanto, ocorreu uma sutil mudança de temáticas pesquisadas: a temática violência parece ter dado lugar aos estudos sobre religião e religiosidade. A classe 2 do primeiro triênio, que tratava sobre as temáticas relacionadas à sociologia econômica, sociedade de consumo, modernidade e

cultura contemporânea, evoluiu dando lugar a outras duas classes no segundo triênio: a classe 2, que tratou das temáticas do desenvolvimento local, rural e territorial e a classe 3, que tratou sobre as temáticas relacionadas à teoria sociológica.

Por meio dos cruzamentos entre as formas e as variáveis consideradas na preparação do *corpus* textual que a AFC realiza, os PPGS podem ser agrupados segundo as classes lexicais resultantes. O Quadro 2 exhibe os programas agrupados pelas classes de cada triênio. Os testes estatísticos que o algoritmo realiza permitem ordenar os PPGS segundo a força da coocorrência dentro da classe, ou seja, quais PPGS são mais 'representativos' em cada classe. A terceira coluna do Quadro 2 mostra o total de instituições atribuídas à classe e a quarta coluna mostra os PPGS com melhor significância para a hipótese de o programa pertencer à classe.

No primeiro triênio tínhamos 42 PPGS listados, no entanto apenas 18 PPGS foram enquadrados nas classes. 3 programas: UFMG-Sociologia Política, UEM e UFAL não tiveram artigos com *Qualis* A1, A2 e B1 neste período (UFAL tinha dois B1, mas eram materiais sem palavras-chave). Outros 21 PPGS não foram agrupados nas classes, pois seus artigos eram muito diversificados, sendo classificados em uma variedade de classes que não permitiu atribuir a cada PPGS uma classe específica.

No segundo triênio tínhamos 52 PPGS e 32 foram agrupados nas classes pela análise lexical das palavras-chave dos artigos publicados. 2 programas não sofreram classificação: FUFPI e UCAM. Estes programas não apresentaram publicação qualificada neste período. Outros 18 PPGS tiveram produção qualificada, mas não foram enquadrados em classes específicas devido à diversidade de temáticas em suas produções (os segmentos de palavras-chave foram classificados, mas cada PPGS não pode ser enquadrado em uma só classe).

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

Quadro 2 – Programas de pós-graduação por classes segundo a produção qualificada

Triênio	Classes	PPGS	PPGS com significância $p < 0,05$ *	Total de PPGS
2007-2009	1	UFPR; UNB; UFG – UNESP/ARAR; UFSM; UFSC; UFPEL	UFPR; UNB; UFG	7
	2	UFSCAR; UFPB/JP; UNISINOS; USP	UFSCAR; UFPB/JP; UNISINOS; USP	4
	3	PUC/MG; UECE-pol. pub. soc.; UECE-plan. Pol. Pub.; UERJ-cien. soc.; UFJF; UFC; UNICAMP-cien. soc.	PUC/MG; UECE-pol. pub. soc.; UECE-plan. Pol. Pub.; UERJ-cien. soc.; UFJF; UFC; UNICAMP-cien. soc.	7
2010-2012	1	UFSC, PUC/MG, PUC/SP, UVV-Sociologia Política, PUC-RIO, UEM, USP, UFU, UVV-Ciências Sociais, UFPEL, UEL, UFMA	UFSC, PUC/MG, PUC/SP, UVV-Sociologia Política, PUC-RIO, UEM, USP, UFU, UVV-Ciências Sociais, UFPEL, UEL, UFMA	12
	2	UFRRJ-Ciências Sociais em Desenvolvimento, Agricultura e Sociedade, UFRGS, UNISINOS, UFG – UFPA-Ciências Sociais	UFRRJ-Ciências Sociais em Desenvolvimento, Agricultura e Sociedade, UFRGS, UNISINOS, UFG	5
	3	UFRJ, UERJ-Sociologia, UFPR, FUFSE, UNICAMP-Sociologia – UFAL, UNB	UFRJ, UERJ-Sociologia, UFPR, FUFSE, UNICAMP-Sociologia	7
	4	UNICAMP-Ciências Sociais, UERJ-Ciências Socioais, UECE- Planejamento e Políticas Públicas, UFRN – UFPB/J.P., UFPE, UNIOESTE, UFC	UNICAMP-Ciências Sociais, UERJ-Ciências Socioais, UECE- Planejamento e Políticas Públicas, UFRN	8

Nota: * estes programas são mais representativos estatisticamente dentro de cada classe ($p < 0,05$)

Fonte: dados processados em *Iramuteq*

A partição de um programa de pós-graduação a uma determinada classe não significa que toda publicação daquele programa seja relacionada àquela subárea temática, visto que os programas possuem autores que publicam em áreas diversas. No entanto, a maior representatividade da produção qualificada de cada programa converge para a classe à qual este foi atribuído. O teste X^2 do *Iramuteq* mostra os programas que apresentam melhor significância dentro de cada classe, são os que estão

listados na quarta coluna do Quadro 2. Estes programas seriam os ‘Líderes de área’, ou seja, os que encabeçaram as classes naquele triênio por apresentarem artigos que trataram de temas semelhantes.

O *Iramuteq* reporta as variáveis mais correlacionadas com os segmentos de texto típicos de cada classe. Com isso, o pesquisador pode realizar uma análise de cunho interpretativo sobre os conteúdos dos artigos científicos dentro de cada classe e agrupá-los por áreas temáticas. Para o estudo foram selecionados os grupos de palavras-chave das publicações dos programas de pós-graduação mais representativos dentro de cada classe. A análise foi concentrada na produção destes programas.

O próximo passo da pesquisa consistiu em realizar uma análise qualitativa de conteúdo da publicação bibliográfica referente aos artigos dos grupos de palavras-chave. O intuito foi o de estabelecer as áreas temáticas manifestas de cada classe, como sintetizamos no Quadro 3.

Quadro 3 – Subáreas temáticas por classes segundo a publicação científica

Triênios	Classes	Subáreas temáticas
2007-2009	1	Política; democracia; partidos políticos; eleições; representação partidária; participação política; políticas públicas; burocracia; teoria sociológica contemporânea; Movimento Sem Terra; reforma agrária; Universidade, ciência e tecnologia.
	2	Sociologia econômica; sociologia financeira; cultura econômica; cultura contemporânea; mercado de trabalho; desigualdade socioeconômica; desigualdade de gênero e raça; classes sociais; mobilidade social movimentos sociais; questões urbanas; pobreza e cidadania.
	3	Violência de gênero; violência doméstica; violência urbana; violência sexual; crime; tráfico de pessoas; prostituição; trabalho doméstico; desigualdade de gênero. Agricultura familiar e Desenvolvimento rural.

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

2010-2012	1	Política e democracia; representação política; conselhos gestores; ciência política; movimentos sociais urbanos; participação política; políticas públicas; burocracia; cidadania; Estado punitivo.
	2	Desenvolvimento rural; agricultura familiar; questões agrárias; reforma agrária; desenvolvimento local; Universidade, ciência e tecnologia; redes de cooperação; justiça; direitos humanos; pobreza; distribuição de renda; direitos sexuais; ativismo social.
	3	Sociologia do trabalho; cultura e sociologia brasileira; pensamento social brasileiro; desigualdade social; ação afirmativa; mobilidade social; teoria sociológica; teoria política; sociologia do conhecimento; sociologia da ciência.
	4	Gênero; sexualidade; movimentos sociais; homossexualidade; etnografia; prostituição; violência doméstica; tráfico; mercado de trabalho; mulheres; saúde; sociologia da religião.

Fonte: análise qualitativa dos *corpora* textuais

Portanto, foi possível estabelecer que a produção mais qualificada da área da sociologia e ciências sociais se encontrava dividida em três grandes subáreas científicas no primeiro triênio e em quatro no segundo. Também ficou evidente a evolução da área em termos de quantidade de programas de pós-graduação, quantidade de produção bibliográfica e diversidade de temas.

3.2 Análise lexical das linhas de pesquisa dos PPGS

A mesma metodologia foi aplicada na análise textual das linhas de pesquisa informadas pela Capes sobre os PPGS durante os dois triênios. O *corpus* textual foi preparado considerando os títulos das linhas de pesquisa e as respectivas descrições que os PPGS informaram dentro de cada ano. As linhas de pesquisa foram codificadas no *corpus* textual com a identificação do PPGS e do ano referente à linha de pesquisa, da seguinte forma:

**** *progr_2 *ano_1

Etnicidade e religiosidade. Estudos sobre

etnicidade e religiosidade a partir de diversos grupos e contextos, observando como tais fenômenos são partilhados e resignificados no mundo contemporâneo.

**** *progr_02 *ano_1

Modernidade, identidades e cultura urbana. Voltada aos estudos sobre os amplos processos de reconfigurações contemporâneas de cultura e identidades, em seus mais variados tipos e manifestações socioespaciais e tendo como foco analítico as pesquisas sobre as sociedades pós_tradicionais.

[...]

A análise lexicográfica básica do *corpus* textual das linhas de pesquisa para o primeiro triênio resultou nos seguintes dados (utilizaram-se os mesmos parâmetros definidos na análise dos artigos):

Número de textos: 583 (segmentos, grupos de Linhas de pesquisa e descrições)
Número de ocorrências: 19818 (palavras)
Número de formas: 1445
Número de lemas (*lemmes*): 1070
Número de formas ativas: 1009 (classes gramaticais)
Número de formas suplementares: 46
Número de formas ativas com frequência > = 3: 846
Média de formas por segmento: 33,99
Número de classes: 4
483 textos classificados de 583 (82,85 %)

Para o triênio 2007-2009, dos 583 segmentos de texto analisados, que equivalem às linhas de pesquisa dos PPGS e suas

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

descrições, 483 foram classificados. O resultado da análise CHD apresentou quatro classes lexicais, como pode ser visualizado na Figura 5.

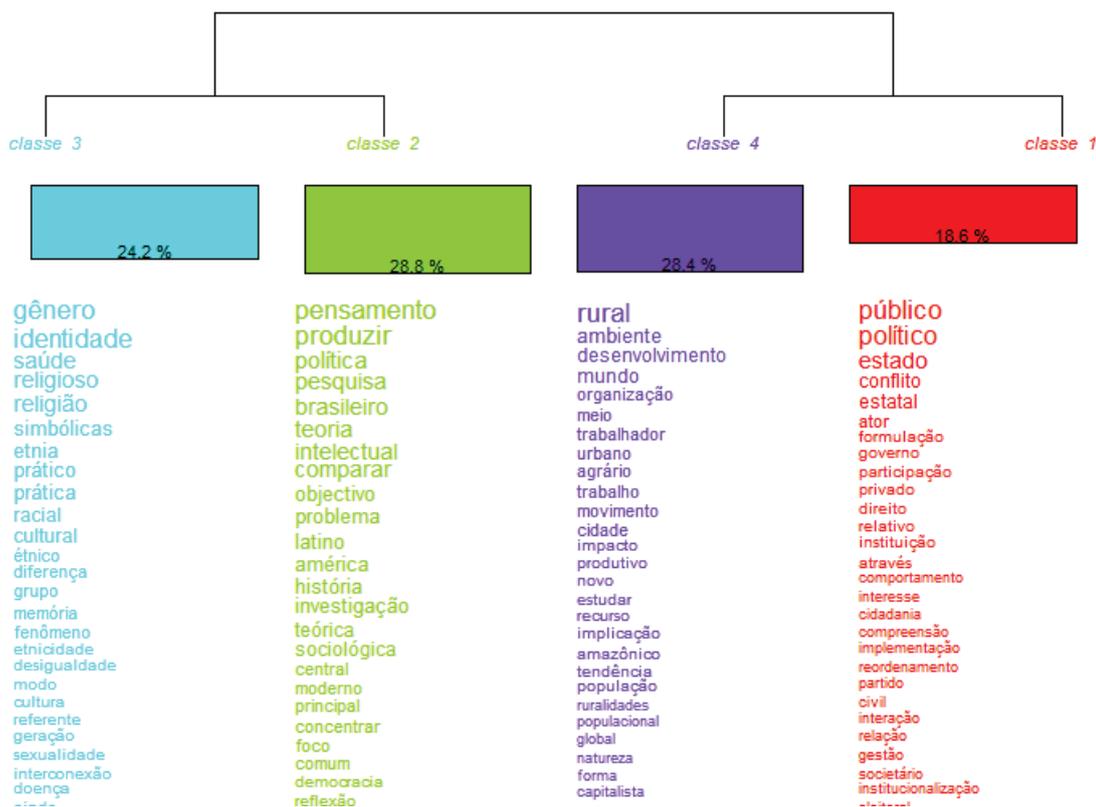


Figura 5 – Dendrograma das classes triênio 2007-2009 segundo linhas de pesquisa

Fonte: corpus processado em Iramuteq

A classe 1 agrupou 90 unidades de contextos elementares de um total de 483 analisados, que equivale a 18,63% do total de linhas de pesquisa consideradas na classificação. A classe 2 agrupou 139, que equivale a 28,78%. A classe 3 agrupou 117, que equivale a 24,22%. Finalmente, a classe 4 agrupou 137 linhas, equivalente a 28,36%. As classes 1 e 4 propõem temáticas de

pesquisa mais próximas, isto é, até certo ponto mais similares, possuem fronteiras teóricas mais próximas. O mesmo ocorre entre as classes 2 e 3.

Para o segundo triênio, a análise das linhas de pesquisa resultou nos seguintes dados:

Número de textos: 626 (grupos de Linhas de pesquisa e descrições)
Número de segmentos de texto: 626 (segmentos considerados)
Número de ocorrências: 21465 (palavras)
Número de formas: 1449
Número de lemas (*lemmes*): 1080
Número de formas ativas: 1017 (classes gramaticais)
Número de formas suplementares: 46
Número de formas ativas com frequência ≥ 3 : 879
Média de formas por segmento: 34,29
Número de classes: 4
544 textos classificados de 626 (86,90 %)

Para o triênio 2010-2012, dos 626 segmentos de texto analisados, que equivalem às linhas de pesquisa dos PPGS e suas descrições, 544 foram classificados. O resultado da análise CHD também apresentou quatro classes lexicais, como pode ser visualizado na Figura 6.

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

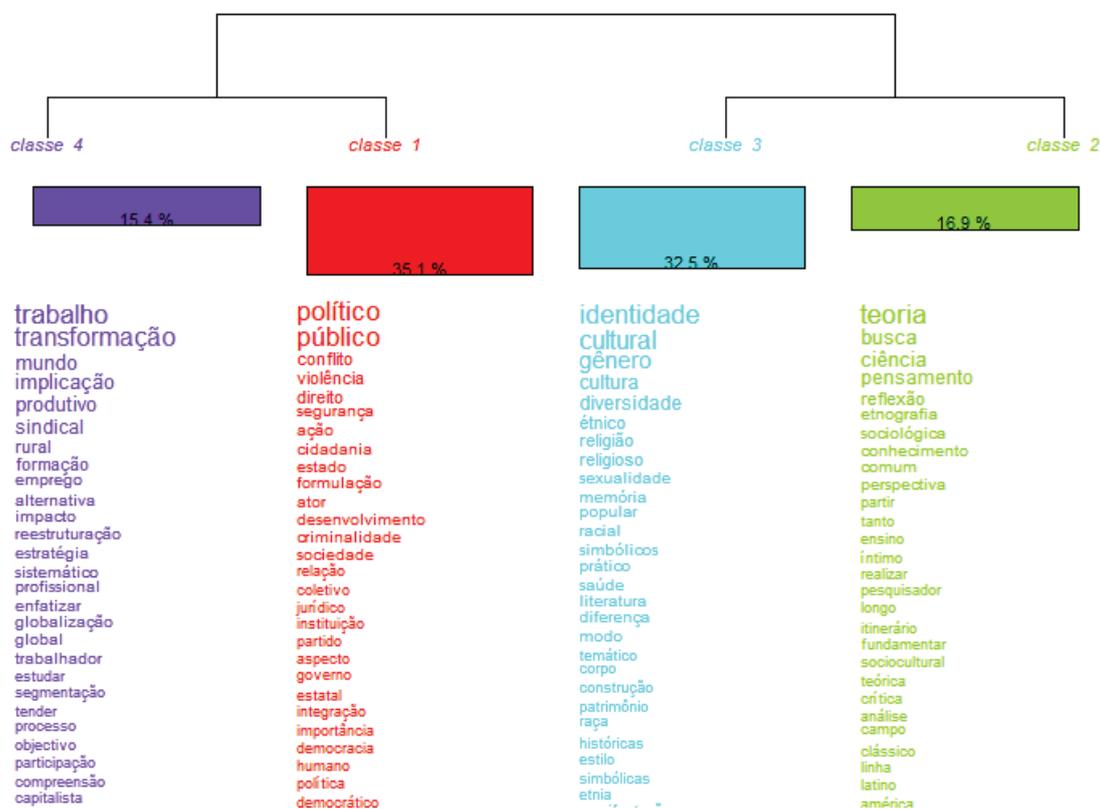


Figura 6 – Dendrograma das classes triênio 2010-2012 segundo linhas de pesquisa

Fonte: corpus processado em Iramuteq

A classe 1 agrupou 191 unidades de contextos elementares de um total de 544 analisados, que equivale a 35,11% do total de segmentos de linhas de pesquisa consideradas na classificação. A classe 2 agrupou 92, que equivale a 16,91%. A classe 3 agrupou 177, que equivale a 32,54%. Finalmente, a classe 4 agrupou 84 linhas, equivalente a 15,44%. De forma análoga ao primeiro triênio, as classes 1 e 4 propõem temáticas de pesquisa mais próximas, isto é, até certo ponto mais similares, e possuem fronteiras teóricas mais próximas. O mesmo ocorre entre as classes 2 e 3.

A proximidade entre as classes nos dois triênios pode ser vista nas figuras 7 e 8, que trazem os resultados da Análise Fatorial de Correspondência. Para o primeiro triênio, o eixo '*facteur 1*' responde por 36,94% do distanciamento entre as classes no tocante ao temas estarem ligados a assuntos de cunho mais teórico (teoria social) à esquerda ou a questões ligadas ao tema desenvolvimento social e econômico, mais à direita. O eixo '*facteur 2*' responde por 33,43% do distanciamento entre as classes quanto a estarem mais próximas aos temas relacionados com política, democracia e público (abaixo) e questões sobre gênero, ruralidade e religião (acima). Um terceiro '*facteur*' corresponde a 29,63%, mas não está representado no eixo cartesiano.

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

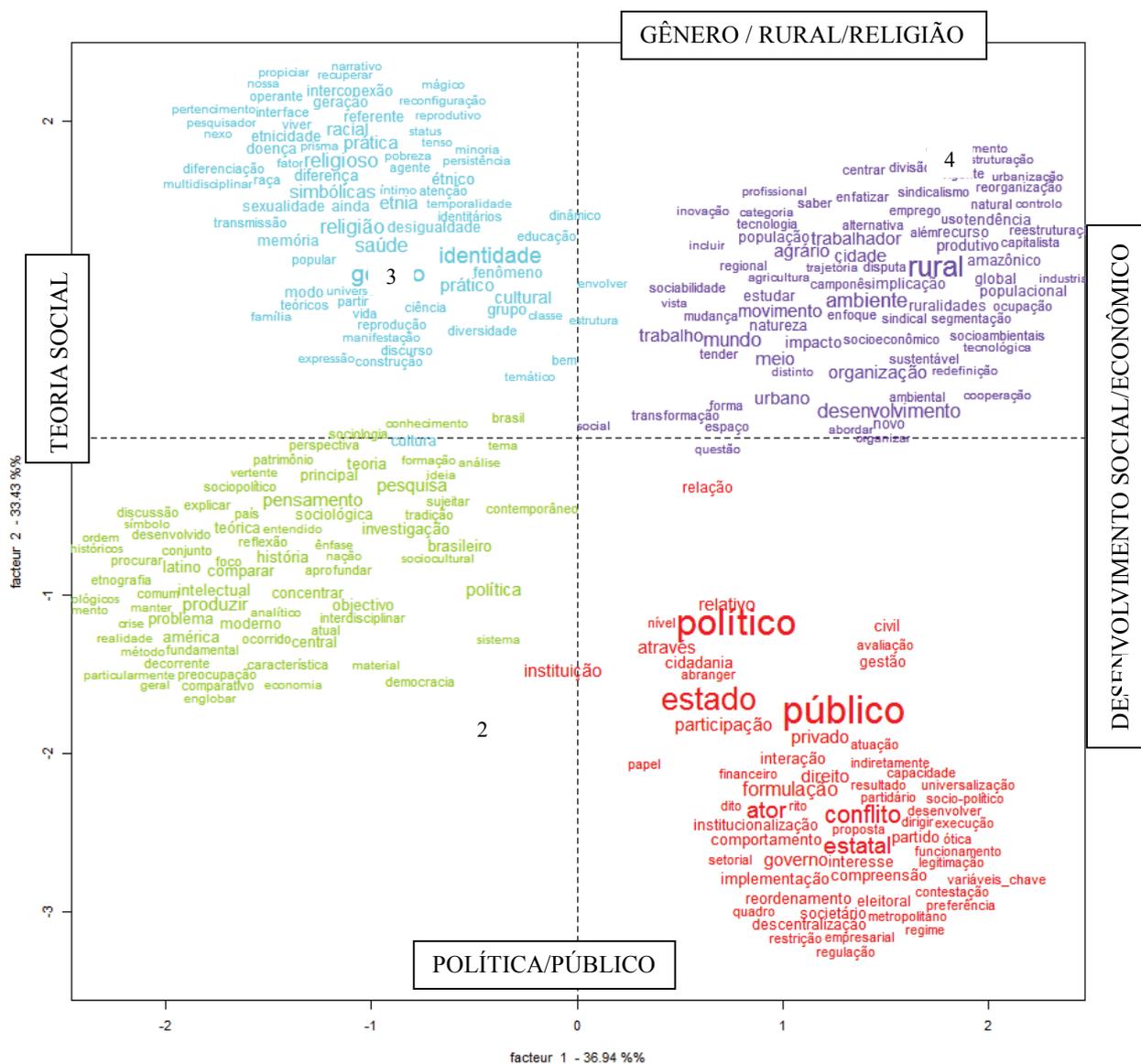


Figura 7 – Quatro classes da Análise Fatorial de Correspondência triênio 2007-2009 (Linhas de pesquisa)

Fonte: corpus processado em Iramuteq

No segundo triênio (Fig. 8) o eixo '*facteur 1*' responde por 37,34% do distanciamento entre as classes no tocante ao temas desenvolvimento social e econômico, mais à esquerda, e temas mais próximos a culturas e identidades, à direita. O eixo '*facteur 2*' responde por 32,91% do distanciamento quanto ao temas estarem mais ligados à teoria e ciência (acima) e questões ligadas a sociologia rural, estudos de gênero e sociologia da religião (abaixo). Um terceiro '*facteur*' corresponde a 29,75%, mas não está representado no eixo cartesiano.

No Quadro 4 exibimos os programas de pós-graduação agrupados segundo as classes lexicais, pelas linhas de pesquisa.

Quadro 4 – PPGS atribuídos por classes no *Iramuteq* segundo as linhas de pesquisa

Triênios	Classes	PPGS	PPGS com significância $p < 0,05$ *	Total de PPGS
2007-2009	1	PUC/RS; UFPB/J.P.; UFRRJ- Ciências Sociais em Desenvolvimento, Agricultura e Sociedade; UFSM; UFMG- sociologia – UNB; UFMG- Sociologia Política	PUC/RS; UFPB/J.P.; UFRRJ- Ciências Sociais em Desenvolvimento, Agricultura e Sociedade; UFSM; UFMG- sociologia	7
	2	UCAM; UNICAMP- Sociologia; USP; UFRJ; UFPR; UNICAMP- Ciências Sociais; PUC/SP – FUFSE	UCAM; UNICAMP- Sociologia; USP; UFRJ; UFPR; UNICAMP- Ciências Sociais; PUC/SP	8
	3	UERJ- Ciências sociais; PUC-RIO; UNESP/ARAR; UFBA – UEM; UFSM; UFJF; UFMG- sociologia; UFCG	UERJ- Ciências sociais; PUC-RIO; UNESP/ARAR; UFBA	9
	4	UFPA- Ciências Sociais; UENF; UFAM; UFSCAR – UNESP/MAR; UFPEL	UFPA- Ciências Sociais; UENF; UFAM; UFSCAR	6
2010-2012	1	PUC/RS; UFMG; UNB; UFG; UFJF; UCAM(IUPERJ) – UFPA- Defesa Social e Mediação de Conflitos; UVV- soc.política; UENF	PUC/RS; UFMG; UNB; UFG; UFJF; UCAM(IUPERJ)	9
	2	FUFSE; UFPE; UNICAMP- Sociologia; UERJ- Sociologia; UFU; UFPEL – UNICAMP- Ciências Sociais; UFPB/J.P.; UFSC	FUFSE; UFPE; UNICAMP- Sociologia;; UERJ- sociologia; UFU; UFPEL	9
	3	UFMA; UFCG; UERJ- Ciências Sociais; UFPA- Ciências Sociais; UFC; UFAL – UNESP/ARAR; UFSM; UEM	UFMA; UFCG; UERJ- Ciências Sociais; UFPA- Ciências Sociais; UFC; UFAL	9
	4	UECE- Políticas Públicas e Sociedade; UFSC; UFSCAR; UFES; UFRGS; UFPEL – PUC/MG; UEL; UECE- Planejamento e Políticas Públicas	UECE- Políticas Públicas e Sociedade; UFSC; UFSCAR; UFES; UFRGS; UFPEL	9

Nota: * estes programas são os mais representativos estatisticamente dentro de cada classe ($p < 0,05$)

Fonte: dados processados em *Iramuteq*

Para o primeiro triênio, 30 PPGS puderam ser atribuídos às classes definidas. Portanto, 12 PPGS não apresentaram

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

conjuntos de linhas de pesquisa suficientemente correlacionados e/ou semelhantes a outros de forma a conseguirem se enquadrar em alguma classe.

No segundo triênio, 36 PPGS puderam ser atribuídos às classes definidas e 16 não puderam ser enquadrados em alguma classe.

Os PPGS agrupados em uma classe são aqueles que possuem linhas de pesquisa similares, isto é, que segundo a coocorrência de termos tratam sobre coisas semelhantes. Ressaltamos que os PPGS podem possuir linhas de pesquisa não somente dentro de uma classe, no entanto, os relacionados na quarta coluna do Quadro 4 (PPGS com significância $p < 0,05$) são aqueles que possuem linhas de pesquisa com maior similitude entre si, podendo ser consideradas representativas para determinar uma subárea dentro da grande área da sociologia.

Para melhor entendimento, selecionamos no Quadro 5 as linhas de pesquisa mais representativas de cada classe. Pela análise de conteúdo destas linhas, foi possível verificar a similitude das linhas dentro de cada classe. Estes PPGS possuem linhas de pesquisa que apresentam pontos em comum, ou seja, linhas que se debruçam sobre temas de pesquisa convergentes.

Quadro 5 – PPGS com linhas de pesquisa mais correlacionadas por classe

(continua)

Triênio	Class e	PPGS	Linhas de pesquisa
2007-2009	1	PUC/RS- Ciências Sociais	Organizações políticas; Organizações sociais
		UFPB/J.P. -Sociologia	Cultura urbana; Trabalho e políticas públicas
		UFRRJ - Ciên. Soc. Desenv., Agric. e Sociedade	Instituições, mercado e regulação; Movimentos sociais; Políticas públicas, estado e atores sociais
		UFSM - Ciências Sociais	Globalização, desenvolvimento e políticas sociais; Instituições e pensamento político
		UFMG- Sociologia	Participação social, políticas públicas e meio ambiente

Velcimiro Inácio Maia

	2	UCAM - Sociologia	Pensamento social no Brasil; Sociedade e criminologia; Sociologia da cultura; Sociologia política; Teoria sociológica e modernidade; Trabalho e sociedade
		UNICAMP - Sociologia	Cidadania e direitos; Estudos de população; Marxismo e sociedade; Pensamento social brasileiro; Sociologia do trabalho
		USP – Sociologia	Cultura, simbolização e representações sociais; Teoria sociológica
		UFRJ – Sociologia e Antropologia	Cultura e política em perspectiva comparada; Sociologia da cultura, simbolismo e linguagem
		UFPR - Sociologia	Cultura e sociabilidades; Instituições e poder; Sociedade e política nas democracias contemporâneas; Teoria, cultura e pensamento social no Brasil América Latina e Europa
		UNICAMP- Ciências Sociais	Cultura e política; Itinerários intelectuais e etnografia do conhecimento; Trabalho, política e sociedade
		PUC/SP - Ciências Sociais	Estado e sistemas sociopolíticos; Materialidade, ideologia e vida cotidiana; Relações internacionais
	3	UERJ - Ciências Sociais	Imagens e perspectivas da subjetividade; Relações étnicas e raciais; Religião e movimentos sociais
		PUC/RIO - Ciências Sociais	Diversidades culturais; Diversidade cultural no Brasil
		UNESP/ARAR - Ciências Sociais	Cultura e ideologia; Cultura, representações simbólicas e pensamento social; Família, relações de gênero, saúde e etnia; Gênero, identidades e etnia; Cultura e pensamento social; Gênero, etnia e saúde
		UFBA - Ciências Sociais	Corporeidade, práticas sociais e narrativas; Relações de gênero, de gerações e raciais em uma sociedade de classes
	4	UFPA - Ciências Sociais	Populações amazônicas, ideias e práticas sociais; Trabalho e sociedade global; Usos sociais dos recursos naturais, madeira, pesca, mineração, recursos hídricos
		UENF - Sociologia Política	Defesa e gestão de recursos naturais e culturais, relações entre população e meio ambiente; Globalização e fragmentação, impactos na sociedade brasileira; Processos globais, sociabilidades e identidades; Território, meio ambiente e processos culturais
		UFAM - Sociologia	Populações amazônicas e formas de representação; Sociedade, meio ambiente e trabalho

Quadro 5 – PPGS com linhas de pesquisa mais correlacionadas por classe

**Produção científica versus linhas de pesquisa dos Programas
de Pós-Graduação em Sociologia e Ciências Sociais no Brasil**

(continuação)

		UFSCAR - Sociologia	Organização social, instituições sociais, organizações formais e informais da sociedade; Tecnologia, desenvolvimento e meio ambiente; Estrutura social, poder e mobilidades; Urbanização, ruralidades, desenvolvimento e sustentabilidade ambiental
2010-2012	1	PUC/RS - Ciências Sociais	Organizações políticas; Organizações sociais; Sociedade e processos políticos
		UFMG - Sociologia	Desigualdades socioeconômicas, organizações e trabalho; Gestão urbana e criminalidade; Participação social, políticas públicas e meio ambiente
		UNB - Sociologia	Educação, ciência e tecnologia; Política, valores e sociedade; Violência, gênero e cidadania
		UFG - Sociologia	Diferença, desigualdade e cidadania; Participação política, instituições e condições sociais da democracia; Violência e criminalidade na sociedade contemporânea
		UFJF - Ciências Sociais	Cultura, democracia e instituições; Políticas públicas e desigualdade social
		UCAM - Sociologia	Direito e sociedade; Juventude, grupos marginalizados e políticas de alteridade; Segurança pública e cidadania
	2	FUFSE - Ciências Sociais	Desenvolvimento socioeconômico e técnica; Política e processos identitários
		UFPE - Sociologia	Educação, trabalho, ciência e tecnologia; Teoria e pensamento social
		UNICAMP - Sociologia	Pensamento social brasileiro; Sociologia do trabalho; Teoria sociológica; Teoria e pensamento sociológico; Trabalho
		UERJ - Sociologia	Filosofia das ciências sociais; Sociologia da cultura; Teoria sociológica e modernidade
		UFU - Ciências Sociais	Cultura, identidades, educação e sociabilidade
		UFPEL - Sociologia	Etnografia e culturas; Teorias sociais e conhecimento
	3	UFMA - Ciências Sociais	Produção social da diferença, minorias nacionais, questões étnicas, raciais e de gênero; Sociabilidades e sistemas simbólicos, cidade, religião e cultura popular
		UFMG - Ciências Sociais	Cultura e identidades
		UERJ - Ciências Sociais	Arte, cultura e política; Imagens e perspectivas da subjetividade; Relações étnicas e raciais; Religião e movimentos sociais
		UFPA - Ciências Sociais	Identidade, etnicidade e gênero, diferenciações e multiplicidades; Populações amazônicas, ideias e práticas sociais; Simbolismo, religião e saúde
		UFC - Sociologia	Cidade, movimentos sociais e práticas culturais; Diversidades culturais, estudos de gênero e processos identitários; Pensamento social, imaginário e religião
		UFAL - Sociologia	Cultura, patrimônio e memória; Trabalho, gênero e saúde

4	UECE - Políticas Públicas e Sociedade	Estado, democracia e participação social
	UFSC - Sociologia Política	Meio ambiente e desenvolvimento urbano e rural; Movimentos sociais, sociedade civil e ecologia política; Mundos do trabalho; Movimentos sociais, participação e democracia
	UFSCAR - Sociologia	Estrutura social, poder e mobilidades
	UFES - Ciências Sociais	Instituições, conflitos e desigualdades
	UFRGS - Sociologia	Sociedade, ruralidade e ambiente; Trabalho e sociedade
	UFPEL - Sociologia	Cidade, estado e esfera pública; Trabalho, organizações e identidade

Fonte: Linhas de pesquisa analisadas em *Iramuteq*

Com foco nos PPGS mais representativos de cada classe, utilizou-se o método qualitativo, por meio da análise de conteúdo das linhas de pesquisa. Assim, foi possível sintetizar as linhas de pesquisa em subáreas de pesquisa, conforme mostramos no Quadro 6.

Quadro 6 – Subáreas temáticas por classes segundo as linhas de pesquisa

Triênio	Classes	Subáreas temáticas
2007-2009	1	Formulação e implementação de políticas públicas; participação social; conflitos sociais; Estado e sociedade civil; instituições políticas e democracia.
	2	Pensamento social no Brasil; criminologia; cultura e política brasileira; teoria sociológica; sociologia do trabalho; sociologia da cultura; sociedade e política contemporâneas.
	3	Gênero; religião; processos identitários; classes sociais; desigualdade social.
	4	Sindicalismo; questão agrária; movimentos sociais; meio ambiente; questões urbanas e rurais; questões amazônicas; desenvolvimento e sustentabilidade ambiental.
2010-2012	1	Desigualdades sociais; desigualdades socioeconômicas; poder nas organizações públicas e privadas; criminologia; violência, criminalidade e políticas públicas; pensamento social brasileiro; segurança pública e direitos humanos; conflitos; democracia; políticas públicas.
	2	Sustentabilidade, ruralidades e urbanidades; sociologia da cultura; teoria sociológica; pensamento político teoria política contemporânea; pensamento social no Brasil; processos socioculturais relacionados ao campo da saúde; processos identitários.
	3	Gênero; desigualdade social; cultura contemporânea; diversidade cultural; pensamento social no Brasil; religião.
	4	Meio ambiente e desenvolvimento; questões agrárias; transformações no mundo do trabalho; sindicalismo, Democracia, participação social e política; desigualdade.

Fonte: análise qualitativa dos das linhas de pesquisa

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

Portanto, pela ótica das propostas dos PPGS na formulação de suas linhas de pesquisa, conseguimos agrupar quatro subáreas temáticas em cada triênio, assim como identificar os PPGS agrupados pela semelhança de suas linhas de pesquisa.

4 Clusterização de PPGS pela produção qualificada versus clusterização de PPGS por linhas de pesquisa

Feito o mapeamento dos agrupamentos de PPGS em torno de temáticas semelhantes em suas linhas de pesquisa e dos agrupamentos de PPGS pelas semelhanças na produção qualificada, podemos partir para a comparação entre as linhas de pesquisa traçadas pelos PPGS e a produção mais qualificada dos mesmos nos dois triênios. O Quadro 7 confronta as subáreas resultantes da análise da produção qualificada com as subáreas resultantes da análise das linhas de pesquisa.

Quadro 7 – Comparativo entre as subáreas segundo linhas de pesquisa e segundo a produção qualificada

Triênios	Classes pela produção bibliográfica	Classes pelas Linhas de pesquisa
2007-2009	1 - Política; democracia; partidos políticos; eleições; representação partidária; participação política; políticas públicas; burocracia; teoria	1 - Formulação e implementação de políticas públicas; participação social; conflitos sociais; Estado e sociedade civil; instituições políticas e

	sociológica contemporânea; Movimento Sem Terra; reforma agrária; Universidade, ciência e tecnologia.	democracia.
	2 - Violência de gênero; violência doméstica; violência urbana; violência sexual; crime; tráfico de pessoas; prostituição; trabalho doméstico; desigualdade de gênero. Agricultura familiar e Desenvolvimento rural.	2 - Pensamento social no Brasil; democracia e instituições; criminologia; política brasileira; teoria sociológica; sociologia do trabalho; sociologia da cultura; sociedade e política contemporâneas.
	3 - Sociologia econômica; sociologia financeira; cultura econômica; cultura contemporânea; mercado de trabalho; desigualdade socioeconômica; desigualdade de gênero e raça; classes sociais; mobilidade social movimentos sociais; questões urbanas; pobreza e cidadania.	3 - Gênero; religião; processos identitários; classes sociais; desigualdade social. 4 - Sindicalismo; questão agrária; movimentos sociais; meio ambiente; questões urbanas e rurais; questões amazônicas; desenvolvimento e sustentabilidade ambiental.
2010-2012	1 - Política e democracia; representação política; conselhos gestores; ciência política; movimentos sociais urbanos; participação política; políticas públicas; burocracia; cidadania; Estado punitivo.	1 - Desigualdades sociais; desigualdades socioeconômicas; poder nas organizações públicas e privadas; criminologia; violência, criminalidade e políticas públicas; pensamento social brasileiro; segurança pública e direitos humanos; conflitos; democracia; políticas públicas.
	2 - Desenvolvimento rural; agricultura familiar; questões agrárias; reforma agrária; desenvolvimento local; Universidade, ciência e tecnologia; redes de cooperação; justiça; direitos humanos; pobreza; distribuição de renda; direitos sexuais; ativismo social.	2 - Sustentabilidades, ruralidades e urbanidades; sociologia da cultura; teoria sociológica; pensamento político; teoria política contemporânea; pensamento social no Brasil; processos socioculturais relacionados ao campo da saúde; processos identitários.
	3 - Sociologia do trabalho; cultura e sociologia brasileira; pensamento social brasileiro; desigualdade social; ação afirmativa; mobilidade social; teoria sociológica; teoria política; sociologia do conhecimento; sociologia da ciência.	3 - Gênero; desigualdade social; cultura contemporânea; diversidade cultural e identidades coletivas; religião.
	4 - Gênero; sexualidade; movimentos sociais; homossexualidade; etnografia; prostituição; violência doméstica; tráfico; mercado de trabalho; mulheres; saúde; sociologia da religião.	4 - Meio ambiente e desenvolvimento; questões agrárias; transformações no mundo do trabalho; sindicalismo, Democracia, participação social e política; desigualdade.

Fonte: Análise qualitativa dos dados

Se compararmos o conjunto total das classes, de acordo com a produção bibliográfica com o total de classes segundo as linhas de pesquisa, podemos perceber que existe certa equivalência entre os temas. No entanto, a comparação classe a

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

classe resulta em poucos pontos de interseção. Procuramos então responder à seguinte questão: A clusterização dos PPGS segundo a produção qualificada estaria alinhada com a clusterização segundo as propostas dos PPGS? A hipótese é de que os agrupamentos de PPGS formados em torno de linhas de pesquisa semelhantes deveriam coincidir com os agrupamentos formados pelas produções bibliográficas semelhantes. Portanto, procedemos à verificação desta hipótese por meio da comparação entre os *clusters*.

No Quadro 8 os agrupamentos da segunda coluna mostram os PPGS que produziram artigos qualificados sobre temáticas semelhantes e os agrupamentos da terceira coluna mostram os PPGS que possuem linhas de pesquisa semelhantes. Consideramos na clusterização apenas os PPGS significantes no teste X^2 da análise fatorial de correspondência (AFC).

Quadro 8 – Comparativo entre agrupamentos de PPGS por produção e por linhas de pesquisa

Triênios	PPGS por classe segundo produção qualificada	PPGS por classe segundo as linhas de pesquisa
2007-2009	UFPR; UNB; UFG	PUC/RS; UFPB/J.P.; UFRRJ- Ciências Sociais em Desenvolvimento, Agricultura e Sociedade; UFSM; UFMG-sociologia
	UFSCAR; UFPB/JP; UNISINOS; USP	UCAM; UNICAMP-Sociologia; USP; UFRJ; UFPR; UNICAMP- Ciências Sociais; PUC/SP
	PUC/MG; UECE-pol. pub. soc.; UECE-plan. Pol. Pub.; UERJ-cien. soc.; UFJF; UFC; UNICAMP-cien. soc.	UERJ-Ciências sociais; PUC-RIO; UNESP/ARAR; UFBA UFPA- Ciências Sociais; UENF; UFAM; UFSCAR
2010-2012	UFSC, PUC/MG, PUC/SP, UVV-Sociologia Política, PUC-RIO, UEM, USP, UFU, UVV-Ciências Sociais, UFPEL, UEL, UFMA	PUC/RS; UFMG; UNB; UFG; UFJF; UCAM(IUPERJ)
	UFRRJ-Ciências Sociais em Desenvolvimento, Agricultura e Sociedade, UFRGS, UNISINOS, UFG	FUFSE; UFPE; UNICAMP-Sociologia; UERJ-Sociologia; UFU; UFPEL
	UFRJ, UERJ-Sociologia, UFPR, FUFSE, UNICAMP-Sociologia	UFMA; UFCEG; UERJ- Ciências Sociais; UFPA-Ciências Sociais; UFC; UFAL
	UNICAMP-Ciências Sociais, UERJ-Ciências Sociais, UECE- Planejamento e Políticas Públicas, UFRN	UECE- Políticas Públicas e Sociedade; UFSC; UFSCAR; UFES; UFRGS; UFPEL

Fonte: dados processados em *Iramuteq*

Podemos verificar se houve ou não convergência por meio da interseção das subáreas (classes). Haverá alinhamento se a interseção for alta. Na comparação do primeiro triênio nos deparamos com um problema: do lado da produção temos três classes e do lado das propostas temos quatro. Mesmo assim, procedemos à comparação por meio da representação das classes e dos PPGS, usando a ‘teoria de conjuntos’, Quadro 9. Por meio de ‘diagramas de Venn’, procuramos verificar qual o grau de interseção entre os conjuntos formados pelas classes de PPGS. Para isso, transformamos as classes em conjuntos numéricos, atribuindo números aos PPGS de acordo com a numeração dada aos PPGS no Quadro 1 no início deste artigo e letras no lugar da numeração das classes. O Quadro 9 mostra as classes segundo os números atribuídos.

Quadro 9 – Classes dos PPGS como conjuntos numéricos

Triênios	PPGS por classe pela produção	PPGS por classe pelas linhas de pesquisa
2007-2009	A = {31; 42; 22 }	D = {5,28,37,40,25}
	B = {39; 28; 49; 50}	E = {7,46,50,34,31,45,6 }
	C = {4; 9; 8; 14; 23; 18; 45}	F = {14,3,43,17 }
		G = {26,12,16,39 }
2010-2012	A = {38,4,6,52,3,11,50,41,51,30,10,24}	E = {5,25,42,22,23,7 }
	B = {37,33,49,22}	F = {2,29,46,13,41,30}
	C = {34,13,31,2,46}	G = {24,19,14,26,18,15}
	D = {45,14,8,35}	H = {9,38,39,20,33,30}

Fonte: elaborado pelo autor

O Quadro 10 resume as interseções entre a clusterização de PPGS nas classes formadas pela produção bibliográfica e nas classes pelas linhas de pesquisa.

Quadro 10 – Interseções entre Classes dos PPGS por produção e Classes por linhas de pesquisa

Triênios	Interseções produção x linhas pesquisa		
2007-	$A \cap D = \{ \}$	$B \cap D = \{28\}$	$C \cap D = \{ \}$

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

2009	AnE = {31} AnF = {} AnG = {}	BnE = {50} BnF = {} BnG = {39}	CnE = {45} CnF = {14} CnG = {}	
2010- 2012	AnE = {} AnF = {41,30} AnG = {24} AnH = {38,30}	BnE = {} BnF = {} BnG = {} BnH = {33}	CnE = {} CnF = {} {13,2,46} CnG = {} CnH = {}	DnE = {} DnF = {} DnG = {14} DnH = {}

Fonte: elaborado pelo autor

Concluimos que há grande divergência entre a clusterização dos PPGS resultante da análise das linhas de pesquisa em relação à clusterização resultante da análise da produção qualificada, visto que os conjuntos de interseção entre as classes se mostraram vazios ou com poucos elementos. Isto mostra que, ao olhar para a área como um todo, não existe alinhamento entre os agrupamentos que se formam pelas linhas de pesquisa com os agrupamentos que se formam segundo a produção qualificada. Isto vai contra o quesito da ficha de avaliação da comissão de área da Capes que considera que a produção acadêmica seja compatível com a área de concentração e as linhas de pesquisa, “a produção técnico-científica deve ser fortemente relacionada à proposta, área(s) de concentração, linhas e projetos de pesquisa” dos PPGS (CAPES, Documento de área de 2013, p.36). No entanto, ressaltamos que aqui estamos observando a área como um todo e não cada PPGS isoladamente, ou seja, se a produção qualificada fosse bem alinhada com as linhas de pesquisa dos programas, os PPGS que se agrupam em torno de temas semelhantes na produção deveriam também se agrupar pela semelhança nas linhas de pesquisa.

5 Considerações finais

Nosso primeiro objetivo foi mapear as subáreas temáticas refletidas pela produção mais qualificada, assim como as subáreas segundo as propostas dos PPGS em suas linhas de pesquisa. Para isso, mostramos que a produção bibliográfica

mais qualificada dos PPGS pôde ser agrupada em três subáreas temáticas no primeiro triênio e em quatro no segundo, enquanto as linhas de pesquisa puderam ser agrupadas em quatro classes em cada triênio.

Quanto ao segundo objetivo traçado, verificar se os agrupamentos de programas segundo a similitude de temas da produção qualificada coincidem com os agrupamentos segundo a similitude das linhas de pesquisa, constatamos que os agrupamentos de PPGS formados pelas temáticas propostas nas linhas de pesquisa não coincidem com os formados pela produção mais qualificada.

Os achados deste estudo não resultaram da análise de convergência entre linhas de pesquisa e produção bibliográfica dentro de cada programa, mas pela clusterização da área da pós-graduação em sociologia segundo estas perspectivas. A comparação das classes obtidas permitiu concluir que os agrupamentos de programas de pós-graduação que tratam sobre temáticas semelhantes não coincidem com os agrupamentos dos programas pelas linhas de pesquisa semelhantes. O que explicaria esta constatação? A pesquisa exploratória realizada nos permite elaborar algumas hipóteses:

a) Os pesquisadores não produzem de forma plenamente alinhada com as linhas de pesquisa de seus programas de pós-graduação, mas sim de acordo com suas expertises (mesmo que as linhas devessem estar alinhadas com as expertises dos docentes). Pode ser também que os autores escrevam sobre temas específicos de forma que não sejam captados na descrição das linhas de pesquisa.

b) Os pesquisadores se debruçam sobre temas atuais de pesquisa (*trending research topics*) ou sobre temas que facilitem o aceite em revistas com *Qualis* mais elevado, mesmo que não tenham grande vinculação com as linhas de pesquisa definidas pelos PPGS.

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

c) Programas novos, que iniciaram cursos de mestrado e/ou doutorado dentro do próprio triênio avaliado, tenderiam a ter uma produção bibliográfica ainda não alinhada com as linhas de pesquisa propostas, visto que tais linhas ainda não foram plenamente consolidadas.

d) A Sociologia ainda é uma área com baixa especialização, ou seja, a sua produção trata de temas muito generalizados.

e) A definição de linhas de pesquisa pode ser clara, mas não é capaz de abarcar o que os pesquisadores realmente fazem, ou seja, criar uma regra não induz as práticas.

A fim de reforçar tais hipóteses, apresentamos os seguintes argumentos:

No primeiro triênio 8 PPGS tiveram pelo menos um curso iniciado no decorrer do período. No segundo triênio foram 11 PPGS novos, sendo que 2 tinham apenas mestrado e iniciaram o curso de doutorado. Esta constatação reforça a hipótese (c).

Se compararmos o número de especialidades de pesquisa na Sociologia com outras áreas do conhecimento, podemos perceber o quanto a Sociologia ainda carece de especialização, reforçando a hipótese (d). Observando a Tabela 2 percebemos que não só a Sociologia, mas também outras áreas das Ciências Humanas como a Antropologia, possuem poucas especialidades. A razão entre quantidade de especialidades e quantidade de PPGS mostra que a área de Sociologia se apresentou aquém de outras áreas do conhecimento que possuem certa afinidade com a Sociologia.

Tabela 2 – Razão entre especialidades do conhecimento segundo Tabela Capes e número de PPGS em 2012

Áreas	N. de PPGS	N. de	Razão N. espec./N.
-------	------------	-------	--------------------

		especialidades	PPGS
Sociologia	52	9	0,17
Antropologia/Arqueologia	26	5	0,19
Ciência Política/Relações Internacionais	32	25	0,78
Economia	55	43	0,78

Fontes: Elaborado pelo autor a partir da Tabela de Áreas do Conhecimento Capes, 2012 e Cadernos de indicadores Capes, 2015.

Desta forma, os resultados aqui apresentados deixam abertos novos caminhos de pesquisa. Estas novas questões poderão ser respondidas por meio de pesquisas qualitativas que possam aprofundar no campo da sociologia e ciências sociais de forma a compreender esta realidade.

Por fim, as limitações desta pesquisa se deram inicialmente pelo corte analítico na publicação bibliográfica, utilizamos somente os artigos mais qualificados, ou seja, com os *Qualis* A1, A2 e B1. Portanto, os achados deste estudo devem ser sempre limitados à produção mais qualificada. Não entraram na análise materiais bibliográficos como livros, teses, dissertações e outros. Desta forma, outras subáreas de pesquisa talvez possam ser encontradas nestes outros substratos ou tipos de publicação, talvez um estudo semelhante a partir das teses e dissertações concluísse em um alinhamento maior com as linhas de pesquisa. Outras limitações são inerentes à metodologia utilizada, advindas das estatísticas utilizadas nos *softwares*, bem como das análises qualitativas realizadas.

Referências:

BARDIN, Laurence. *Análise de Conteúdo*. Lisboa: Edições 70, 1977.

CAMARGO, Brígido Vizeu. Alceste: um programa informático de análise quantitativa de dados textuais. In: MOREIRA, et al. (Org.). *Perspectivas teórico-metodológicas em*

Produção científica versus linhas de pesquisa dos Programas de Pós-Graduação em Sociologia e Ciências Sociais no Brasil

representações sociais. João Pessoa: Editora Universitária UFPB, 2005. p. 511-539.

CAMARGO, Brigido Vizeu; JUSTO, Ana Maria. IRAMUTEQ: um software gratuito para análise de dados textuais. *Temas em Psicologia*, Ribeirão Preto, v. 21, n. 2, p. 513-518, 2013.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. CAPES. *Tabela de Áreas do Conhecimento Capes*, 2012. Disponível em: <<http://www.capes.gov.br/avaliacao/instrumentos-de-apoio/tabela-de-areas-do-conhecimento-avaliacao>>. Acesso em: 23 nov. 2015.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. CAPES. *Cadernos de Indicadores*. 2015. Disponível em:

<<http://conteudoweb.capes.gov.br/conteudoweb/CadernoAvaliacaoServlet>>. Acesso em: 25 nov. 2015.

COORDENAÇÃO DE APERFEIÇOAMENTO DO PESSOAL DE ENSINO SUPERIOR. CAPES. *Documento de área 2013*. 2013. Disponível em:

<http://www.capes.gov.br/images/stories/download/avaliacaotrienal/Docs_de_area/Sociologia_doc_area_e_comiss%C3%A3o_21out.pdf>. Acesso em: 20 out. 2015.

JUSTO, Ana Maria; CAMARGO, Brigido Vizeu. Estudos qualitativos e o uso de *softwares* para análises lexicais. In: NOVIKOFF, C.; SANTOS, S. R. M. & MITHIDIERI, O. B. (Orgs.). *Caderno de artigos: X SIAT & II Serpro*. Duque de Caxias: Universidade do Grande Rio "Professor José de Souza Herdy" – UNIGRANRIO, 2014. p. 37-54. Disponível em: <<https://lageres.wordpress.com/>>. Acesso em: 20 out 2015.

- MACHADO, R. das N. Produtividade dos autores em ceratocone: estudo longitudinal em artigos indexados no SciELO Brasil (2001-2010). *RDBCI*, v. 11, n. 1, p. 1-20, 2013. Disponível em: <<http://www.sbu.unicamp.br/seer/ojs/index.php/rbci/articloe/view/558/pdf>>. Acesso em: 16 jul. 2015.
- NASCIMENTO, Adriano Roberto Afonso do; MENANDRO, Paulo Rogério Meira. Análise lexical e análise de conteúdo: uma proposta de utilização conjugada. *Estudos e pesquisas em psicologia*, Rio de Janeiro, UERJ, ano 6, n. 2, 2º sem. 2006.
- REINERT, M. Alceste, une méthodologie d'analyse des données textuelles et une application: Aurelia de Gerard de Nerval. *Bulletin de Methodologie Sociologique*, v.26, p.24-54, 1990.
- REINERT, M. Alceste: Analyse de données textuelles. *Manuel d'utilisateur*. Toulouse: IMAGE, 1998.