

**CENTRO DE ESTATÍSTICA APLICADA – CEA – USP**  
**RELATÓRIO DE CONSULTA**

**TÍTULO DO PROJETO:** “A Sintaxe Verbal no Baixo Alemão dos Mennonitas nas Américas”.

**PESQUISADOR:** Göz Kaufmann.

**INSTITUIÇÃO:** Faculdade de Filosofia, Letras e Ciências Humanas

**FINALIDADE DO PROJETO:** Publicação.

**PARTICIPANTES DA ENTREVISTA:** Göz Kaufmann.  
Carlos Alberto Bragança Pereira  
Cláudia Monteiro Peixoto  
Fernando Henrique Ferraz Pereira da Rosa  
Vagner Aparecido Pedro Júnior

**DATA:** 30/09/2005

**FINALIDADE DA CONSULTA:** Auxílio na descrição e modelagem dos dados.

**RELATÓRIO ELABORADO POR:** Fernando Henrique Ferraz Pereira da Rosa  
Vagner Aparecido Pedro Júnior

## 1. Introdução

A história dos Mennonitas é antiga e tem suas raízes na Suíça e Alemanha, da época da reforma de Lutero. Por serem anabatistas (sendo batizados apenas quando adultos), foram perseguidos numa época de pouca tolerância religiosa. Assim espalharam-se pela Europa, sendo que alguns grupos estabeleceram-se na região da atual Holanda, onde o sacerdote e líder Menno Simons os redenominou de "Mennonitas".

No início da 1ª Guerra Mundial, os Mennonitas chegaram a somar cerca de 120.000 pessoas, todavia, com a revolução de 1917 e a tomada do poder pelos comunistas, os Mennonitas perderam privilégios, tendo sofrido desapropriações e perseguições. Em 1929, permaneciam na Holanda apenas cerca de 6.000 mennonitas, que então migraram para a Alemanha e a China. Mais tarde, através da imigração, a Igreja Irmãos Menonitas expandiu-se para as Américas do Norte e do Sul.

Nesse contexto, foi realizado um estudo que pretende identificar em diferentes colônias a forma como ocorreu a evolução da língua germânica. Essas colônias são: Chihuahua no México, Menno e Fernhein no Paraguai, Santa Cruz na Bolívia, Texas nos Estados Unidos e Rio Grande do Sul no Brasil.

## 2. Exploração

Para fazer a identificação da evolução da língua germânica, foram feitas entrevistas Sociolinguísticas com 325 Mennonitas nas 6 colônias. A cada entrevistado foram apresentadas 46 orações na língua local de sua colônia (espanhol, inglês ou português) para que ele as traduzisse para o alemão, sempre oralmente, com a intenção de obter respostas espontâneas.

Cada oração traduzida pode ser classificada em três variantes diferentes da língua alemã: NR (*Non-Raising Variant*), VPR (*Verb Projection Rasing*) e VR (*Verb Raising*). Essa classificação será a variável resposta e é determinada através da colocação verbal na tradução feita pelos entrevistados. Foram utilizadas frases com duas e três formas verbais. A seguir um exemplo de como identificá-las, para duas e três formas verbais.

### Duas Formas Verbais

- *Non-Raising Variant* que demonstra o alemão padrão;

Ex: *Sei weit, daut sei vondaag IHRA MAMA helpe soll.*

- *Verb Projection Raising*, de origem flamenca;

*Ex: Sei weit, daut sei vondaag **soll** IHRA MAMA helpe.*

- *Verb Raising* de origem Holandêsa.

*Ex: Sei weit, daut sei vondaag IHRA MAMA **soll** helpe.*

Sendo a tradução dessa frase em português:

Ela sabe que hoje **deve** ajudar à mãe dela.

### Três Formas Verbais

- *Non-Raising Variant:*

*Ex: Sei weit, daut sei gesteren IHRA MAMA helpe **hat** **sollt**.*

- *Verb Projection Raising:*

*Ex: Sei weit, daut sei gesteren **hat** IHRA MAMA helpe **sollt**.*

*Ex: Sei weit, daut sei gesteren IHRA MAMA **hat** helpe **sollt**.*

- *Verb Raising:*

*Ex: Sei weit, daut sei gesteren **hat** **sollt** IHRA MAMA helpe.*

*Ex: Sei weit, daut sei gesteren **hat** IHRA MAMA **sollt** helpe.*

*Ex: Sei weit, daut sei gesteren IHRA MAMA **hat** **sollt** helpe.*

Sendo a tradução dessa frase em português:

Ela sabe que ontem **devia ter** ajudado à mãe dela.

Além disso, foram medidas as seguintes variáveis independentes, para cada entrevistado (ou falante) e frase:

- Idade (em anos ou categorizada em 3 níveis) do falante.
- Sexo (masculino e feminino) do falante.
- Variáveis de caracterização da frase:

- Tipo de Oração (Causal – Complemento – Condicional – Relativo);
- Tipo do verbo finito (Modal/ *woare* – *han* – *dun(e)*);
- Marcação morfológica do artigo definido (*daut* – *dem* – *den* – *de/die*).

A intenção do pesquisador é identificar como se comporta a escolha da variante verbal (variável resposta) de acordo com o tipo de oração, tipo do verbo, marcação morfológica do artigo definido, sexo e idade do falante. Uma potencial complicação é a dependência entre frases traduzidas pelo menos indivíduo. Dependendo do grau dessa dependência, isso deve ser levado em conta nos procedimentos de análise propostos.

### 3. Sugestão do CEA

Assumindo que as variáveis independentes medidas (caracterização da frase, sexo e idade) expliquem a variação individual, ou tenham um efeito na resposta comparativamente maior que o da variação individual, podemos considerar cada frase como uma unidade amostral independente. Nesse caso, considerando a idade categorizada, os dados se resumem a uma tabela de contingência com 3 níveis de resposta e  $3 \times 2 \times 3 \times 3 \times 3 = 162$  subpopulações. Nessa situação, podemos propor o ajuste de um modelo de contagem de frequências de Poisson ou o modelo saturado multinomial, conforme descrito em Agresti (2002) e Venables e Ripley (2002).

Outra possibilidade, também descrita em Agresti (2002), é o ajuste de um modelo de regressão logística politômico, que relaciona particulares níveis das variáveis independentes à probabilidade de cada variante (NR, VPR, VR).

Caso não seja possível considerar as frases independentes (quando as variáveis independentes por si não são o bastante para explicar a dependência entre as frases), deve-se considerar outras estratégias de análise. Pode-se separar por frase e fazer uma análise para cada tipo de frase, ou atribuir escores a cada frase, fazendo a análise então a partir desses escores.

### 4. Conclusão

A análise desse conjunto de dados requer um planejamento cuidadoso e informações como o grau de dependência das frases entre indivíduos precisam ser avaliadas antes de ser possível decidir qual a abordagem de análise mais vantajosa.

Caso haja interesse do pesquisador, sugere-se a submissão desse projeto no processo de triagem do CEA para a realização de um estudo detalhado, utilizando partes do conjunto de dados.

## Referências

AGRESTI, A. (2002). **Categorical Data Analysis**. 2.ed. New Jersey: John Wiley & Sons. 526p.

VENABLES, W.N. e RIPLEY, B. D. (2002). **Modern Applied Statistics with S**. 4.ed. New York: Springer. 495p.