



# DOCUMENTAÇÃO LINGÜÍSTICA

## POR QUE, PARA QUEM, COMO

Bruna Franchetto & Mara Santos (UFRJ)

Kristine Stenzel (UFRJ)

Sebastian Drude (MPEG)

Março 2009

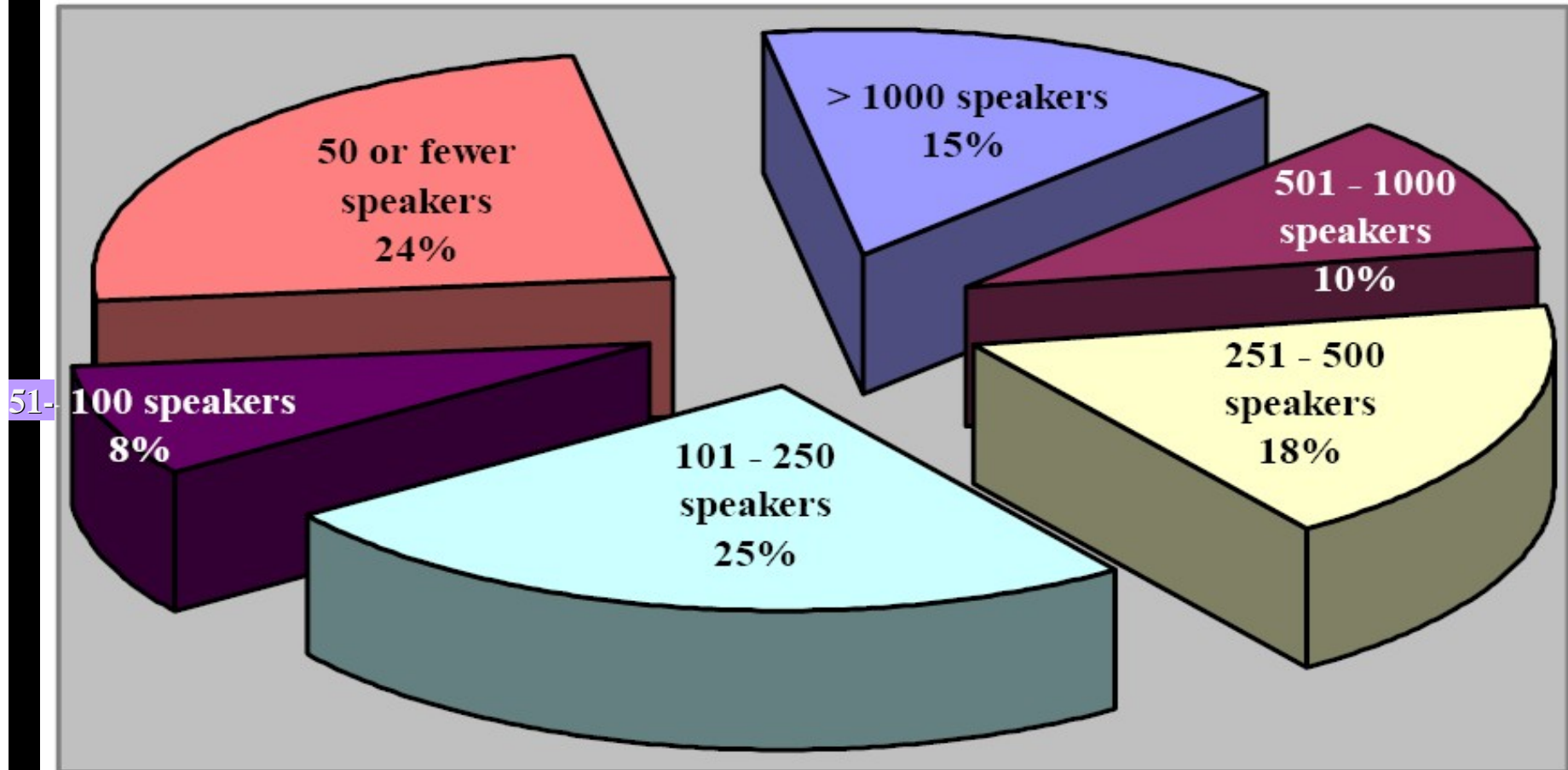
# Diversidade Lingüística

- no mundo existem ca. 5.000–6.000 línguas
- não há um número exato, porque a diferenciação entre língua e dialeto é variável
- as línguas pertencem a centenas de famílias lingüísticas e há muitas línguas isoladas
- a distribuição é muito heterogênea
- o tamanho das comunidades de falantes varia muito

- nas Américas há ca. 1.000 línguas
- no Brasil são ca. 160 línguas indígenas (2 troncos, 40 famílias, 10 isoladas) (Moore, 2008)
- em 1500 eram possivelmente 600 línguas
- o tamanho médio é ca. 270 falantes por língua
- pequeno número de falantes (200) é típico para as línguas das terras baixas tropicais (cf. Pápua, África Equatorial)
- há várias línguas de uso restrito ou inativas
- muitas línguas já não são transmitidas para as novas gerações

# Distribuição das línguas indígenas no Brasil por número de falantes

## Indigenous Language Groups in Brazil



Source: Moore and Gabas Jr. (in press)

# Perda da diversidade lingüística

- Perda populacional: genocídio (processo rápido e devastador) e migração (processo rápido ou gradual)
- Perda de falantes e da fala: (processo gradual de abandono e ‘troca lingüística’)
  - Pressões externas: exploração e integração de áreas remotas na escala nacional e global, acompanhado por marginalização dos grupos minoritários, proibição ou desvalorização de suas práticas e símbolos culturais (inclusive o uso das línguas)
  - Pressões internas: internalização do discurso dominante - perda de prestígio da língua entre os próprios falantes, LI vista como inferior, um empecilho à ‘integração’; **decisão** de abandonar a língua!
- falta de escolaridade em línguas indígenas; língua nacional domina as mídias (televisão, rádio, internet, imprensa, etc.).

# Línguas e Culturas em Perigo

Documento UNESCO (2003) *Language Vitality and Endangerment* - critérios para avaliar a vitalidade de uma língua:

1. Número absoluto de falantes
2. Proporção de falantes na população
3. Transmissão inter-geracional
4. Mudanças nos domínios de uso da LI
5. Resposta a novos domínios de uso, mídias
6. Existência de materiais pedagógicos, de alfabetização e de pesquisa em LI

# Documentação Lingüística

- Pelo menos desde 1992 o problema é percebido e discutido entre os lingüistas
- No mínimo é possível **documentar**, inclusive para uma possível revitalização
- Um novo ramo: a **lingüística documental**
- **‘Documentação’** ganha um novo sentido, complementar à **descrição** (p. ex., uma gramática e um dicionário)
- Objetivo: A criação de **acervos sustentáveis** digitais que registram o **uso** da língua

## Surgem programas de documentação

- A partir dos anos 90 surgem programas de documentação de línguas ameaçadas
- Alguns dos principais, hoje:
  - DOBES (VWS, Alemanha)
  - ELDP (SOAS, London)
  - EMELD (LinguistList)
  - AILLA (U. Texas)
  - PARADISEC (Austrália)
  - LACITO (França)
  - LDA (LDC, U. Pennsylvania)
  - ELF (Yale U.)

# Documentação Lingüística no Brasil

- No Brasil, a Documentação Lingüística se estabeleceu nos últimos 8 anos em cooperação com programas internacionais: DOBES (Alemanha) e ELDP (Inglaterra)
- 2007: primeiros acordos para transferência de tecnologia assinados entre FUNAI e Ministério de Ciência e Tecnologia (MCT) e o Instituto Max Planck de Psicolingüística (MPI, Programa DOBES)
- A tecnologia para armazenar dados foi desenvolvida e está instalada em servidores no Museu do Índio/FUNAI e Museu Paraense Emílio Goeldi/MCT

# Projetos DOBES no Brasil

- **Kuikuro** (Alto Xingu, Karib, Bruna Franchetto, Museu Nacional / UFRJ)
- **Trumai** (Alto Xingu, isolada, Raquel Guirardello-Damian, MPI Nijmegen & Museu Goeldi)
- **Awetí** (Alto Xingu, Tupí, Sebastian Drude, Freie Univ. Berlin & Museu Goeldi)
- **Mawé** (Tupí), **Bakairí**, **Kashuyana** (Karib)  
(Sérgio Meira, Leiden & Museu Goeldi)
- **Kaxinawa** (Pano, E. Camargo, Paris X & UNICAMP)
- **Aché** (TG, E, Roessler, ULB & UNICAMP)

# Projetos ELDP/SOAS no Brasil

- **Apurinã** (Aruak, Sidney Facundes, UFPA)
- **4 lgs. Tupí** (D. Moore etc, Museu P.E. Goeldi)
- **Lgs. Chaco** (V. Grondona & F. Sândalo, UNICAMP)
- **Ofayê** (Macro-Jê, Eduardo Ribeiro, UFGO)
- **Karo** (Tupí, Nilson Gabas Jr., M.P.E. Goeldi)
- **Enawene-Nawe** (Aruak, Ubiray Rezende, Museu Nacional / UFRJ)
- **Waikhana + Wanano** (Tukano Oriental, K. Stenzel, Museu Nacional / UFRJ)
- **Pareci-Haliti** (Arawak, G. Romling da Silva, UFRJ)

# DOCUMENTAÇÃO LINGÜÍSTICA

## O que é?

# DOCUMENTAÇÃO LINGÜÍSTICA 'CLÁSSICA'

## Objetivos:

- Elaborar uma Gramática e um Léxico
- Constituir um *corpus* de textos

manuscritos / impressos / digitais

em formatos: word, txt, etc.

# DOCUMENTAÇÃO LINGÜÍSTICA, ONTEM E HOJE:

Objetivos:

- permitir a recuperação do conhecimento de uma língua mesmo quando todas as outras fontes estão perdidas
- gerar recursos para a manutenção e/ou aprendizagem de uma língua

# DOCUMENTAÇÃO LINGÜÍSTICA HOJE

## Objetivos:

- coletar/criar materiais documentais áudio, vídeo, gráficos, mapas e textos incluindo o uso de uma língua numa variedade de contextos sociais e culturais
- coletar, gravar, arquivar e analisar
- criar materiais de boa qualidade para embasar a descrição de fenômenos lingüísticos

# A DOCUMENTAÇÃO PODE PRODUZIR MATERIAIS EM DIVERSOS TIPOS DE MÍDIA

- vídeo
- áudio
- imagens (mapas, iconografia, fotos, vídeos, desenhos, etc.)
- escritos (notas de campo, transcrições, artigos, análises, etc.)
- metadata (dados estruturados sobre cada material, em forma escrita)

# MATERIAIS QUE PODEM SER PRODUZIDOS POR UMA DOCUMENTAÇÃO LINGÜÍSTICA

- informações etnográficas e sociológicas relevantes
- dicionário
- materiais pedagógicos
- gramática e descrição lingüística
- um bom corpus de dados

# UM BOM CORPUS DE DADOS DEVE TER E DEVE SER:

- diversidade (situações, participantes, canais, gêneros de fala, variedades)
- amplo
- em construção
- transparente
- preservável (metadata, dados arquivados, dados gravados e armazenados digitalmente)
- transportável (tem que poder migrar para novas tecnologias)
- ético

# O que é Documentação Linguística, hoje

- Uma documentação moderna contém um corpus de **dados primários**, gravações em áudio e/ou vídeo de uma grande variedade de domínios e situações de uso
- Os dados devem ser organizados e rotulados conforme seu conteúdo e data e acompanhados por **anotações**
- O mínimo de anotações é uma **transcrição** e uma **tradução**
- Uma documentação deve ser **digital**, em formatos **estandardizados e permanentes, acessível** ao grande público, p.ex. via internet, através de **Meta-Data**

# Dados Primários

## Lingüísticos e Culturais

1. Documentação hoje : a construção de acervos multimídia
2. o que documentar? : etnografia e documentação
3. a participação da comunidade de falantes
4. como documentar? : equipamentos e tecnologia
5. gerenciamento de energia no campo
6. digitalização/captura, conversão, edição e armazenamento de dados primários

# 1. A Construção de Acervos Multimídia

A relação entre **descrição lingüística** e **documentação lingüística** - dois campos de atuação *complementares*, porém *diferentes* em termos de:

- atividades primárias
- métodos e procedimentos
- resultados imediatos

**Descrição lingüística** - investigação aprofundada de um sistema lingüístico, visando a produção de uma gramática, um dicionário e uma coletânea de textos dirigida basicamente a uma audiência de lingüistas

atividades primárias: análise, acompanhado por definições de termos e conceitos e justificativas de validade

métodos e procedimentos: elicitación dirigida, análise de textos de um corpus restrito (consultores, local, tempo)

resultados imediatos: conjunto de afirmações descritivas (sobre fenômenos sincrônicos: fonéticos, fonológicos, morfossintáticos, semânticos), ilustrados com exemplos

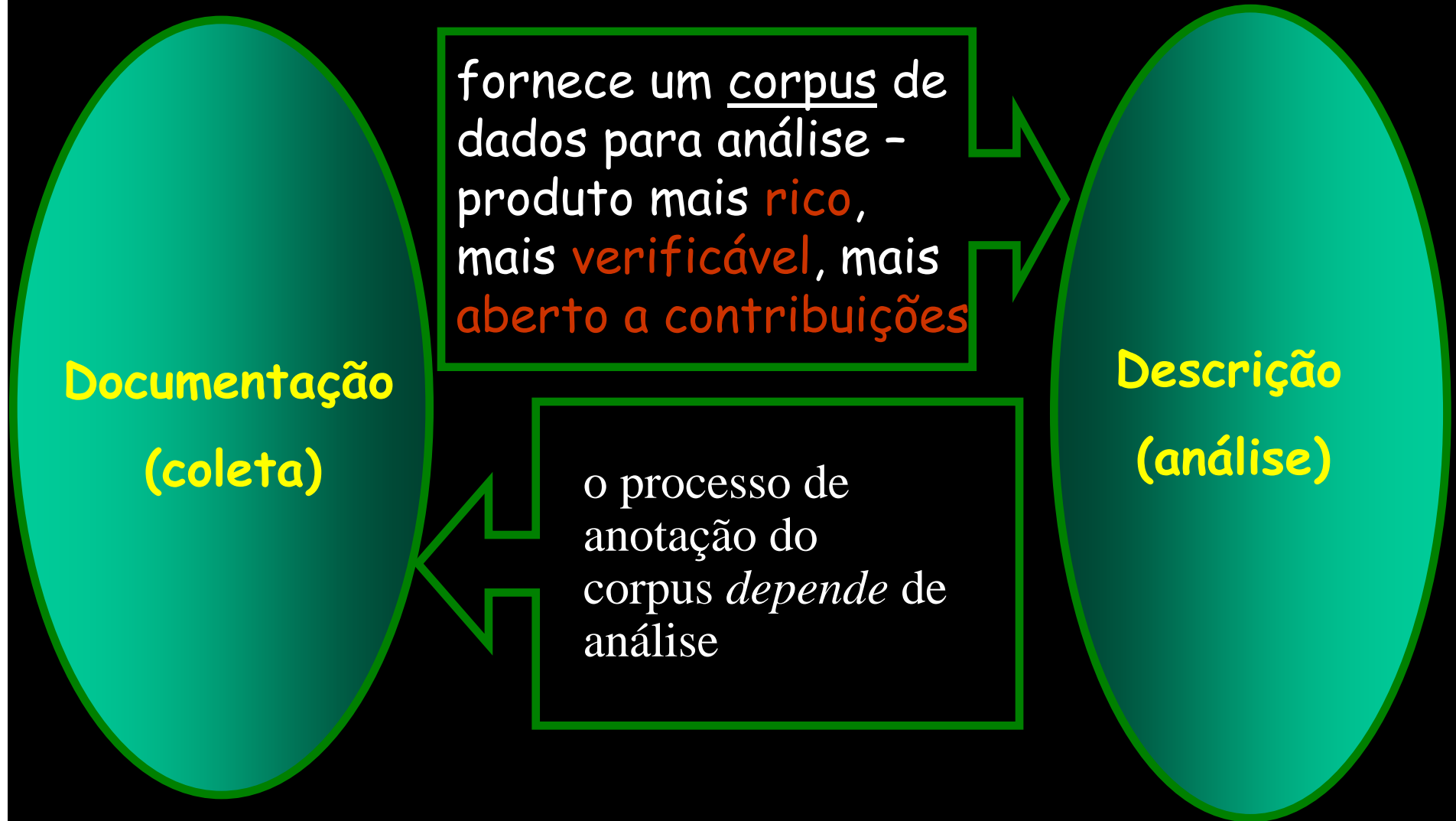
**Documentação lingüística** – registro sistemático e amplo de exemplos de uso de uma língua em contextos culturais naturais e apropriados, os mais variados possível, visando a constituição de um corpus anotado

atividades primárias: coleta de dados, acompanhado por anotação sistemática, arquivamento

métodos e procedimentos: gravações, transcrição, tradução, arquivamento (preservação)

resultados imediatos: corpus de ocorrências de fala, com anotações metalingüísticas (falantes, pesquisador(es), situação, conteúdo, etc.)

# Como a descrição e a documentação se complementam e se fortalecem (Woodbury 2003)



# As 'novidades' da documentação

- Foco em dados primários – coleta ampla – ‘corpus’ disponibilizado para um conjunto variado de usuários;
- Preocupação com a responsabilidade/verificabilidade – acesso aos dados primários facilita a avaliação de análises;
- Preocupação com a manutenção e preservação dos dados a longo prazo – garantir a disponibilidade para usuários no futuro;
- Trabalho em equipes interdisciplinares – não é restrita somente à comunidade dos lingüistas;
- Participação da comunidade de falantes – trabalho colaborativo da comunidade como produtores e usuários de materiais e como co-pesquisadores.

## 2. Etnografia e Documentação: a interdisciplinariedade

(Franchetto 2006), (Widlok 2005)

Para o etnólogo, as línguas representam o acesso a ‘categorias nativas’ e, portanto, ao entendimento de cosmologias, estruturas sociais, rituais, relações entre mundos humanos e não-humanos, etc.

A DL é uma boa ferramenta para a etnografia:

- registra gêneros variados de fala
- fornece descrições de contextos de eventos de fala
- ajuda na identificação de termos e expressões significativas

## A documentação colabora com:

- biologia (botânica, zoologia, ecologia)
- etnomusicologia
- arqueologia
- educação
- outras áreas da lingüística: aquisição de linguagem, análise de conversação, línguas de sinais, contato e mudança, bi-/multilingüismo, sócio-lingüística, tipologia, entre outros.

### 3. A Participação da Comunidade

- Mudança: da pesquisa ‘sobre (uma língua, um povo)’ à pesquisa ‘com’ (um povo, os falantes de uma língua)
- o que é ‘comunidade’? quem participa? (falantes, meio-falantes, líderes, professores e alunos...)
- Vantagens da pesquisa ‘com’:  
+ produtivo; + representativo; de efeito + duradouro
- Dificuldades e limitações:  
políticas internas e externas; abertura/flexibilidade /mudanças de rumo; treinamento; tempo

# Bibliografia

- Drude, Sebastian (2006) “A metodologia DOBES de documentação lingüística e o formato de anotação de textos”. *Estudos Lingüísticos* 35. 27–51. [<http://gel.org.br/4publica-estudos-2006/sistema06/sd.pdf>].
- Franchetto, Bruna (2005). “Línguas ameaçadas e línguas como patrimônio imaterial: duas idéias em discussão”. *Revista do Patrimônio Histórico e Artístico Nacional, no 32: Patrimônio imaterial e biodiversidade*. Manuela Carneiro da Cunha (org.). Brasília IPHAN/Minc. (182-205).
- Franchetto, Bruna (2006). “Ethnography in language documentation” in Gippert, Jost; Himmelmann, Nikolaus P.; Mosel, Ulrike (orgs.) (2006) *Essentials of Language Documentation*. Berlin, New York: Walter de Gruyter (p.183-212)
- Gippert, Jost; Himmelmann, Nikolaus P.; Mosel, Ulrike (orgs.) (2006) *Essentials of Language Documentation*. Berlin, New York: Walter de Gruyter.
- Grenoble, Lenore A. & Whaley, Lindsay J. (1988). *Endangered Languages*. Cambridge: Cambridge University Press.
- Himmelmann, Nikolaus P. (2002) “Documentary and Descriptive Linguistics (full version). IN O. /sakiyama & F. Endo (eds) *Lectures on Endangered Languages 5*. Kyoto: Endangered Languages of the Pacific Rim.
- Himmelmann, Nikolaus P. (2006) “Language documentation: what is it and what is it good for?” in Gippert, Jost; Himmelmann, Nikolaus P.; Mosel, Ulrike (orgs.) (2006) *Essentials of Language Documentation*. Berlin, New York: Walter de Gruyter (p.1-30)
- Leite, Yonne & Franchetto, Bruna. (2006). “500 anos de línguas indígenas no Brasil.” In Cardoso, Suzana A.M., Mota, Jacyra A., Mattos e Silva, R.M. (orgs) *Quinhentos Anos de História Lingüística no Brasil*. Salvador: Sec. da Cultura e Turismo do Estado da Bahia (p. 15-61)
- Moore, Denny. (2005). “Brazil: Language Situation” In *Encyclopedia of Languages and Linguistics*, 2nd. edition. Keith Brown (ed.) V. 2 Amsterdam: Elsevier. (p. 117-127).
- Moore, Denny (2006): “Endangered languages of Lowland tropical South America”. In: M. Brenzinger (org.). *Endangered Languages and Language Documentation*.
- Moore, Denny and Gabas Jr., Nilson. (2006). “O futuro das línguas indígenas brasileiras.” In *Amazônia 500 Anos*. Louis Forline, Ima Vieira & Rui Murrieta (orgs). Belém: Museu Paranaense Emilio Goeldi. (p. 333-354)
- Nettle, Daniel & Romaine, Suzanne. (2000). *Vanishing Voices: The Extinction of the World's Languages*. Oxford: Oxford University Press.
- Ricardo, Carlos Alberto & Ricardo, Fany. (2006). *Povos indígenas no Brasil 2001-2005*. São Paulo: Instituto Socioambiental.
- Rodrigues, Aryon Dall'Igna. (1993). “Línguas Indígenas: 500 anos de descobertas e perdas.” *Ciência Hoje*, 16 (p. 20-26).
- Rodrigues, Aryon Dall'Igna. (2005). “Sobre as línguas indígenas e sua pesquisa no Brasil.” *Ciência e Cultura*, 57. (p. 35-38).
- Stenzel, Kristine. (2006) “Línguas indígenas e tradições orais na Amazônia Brasileira”. In: *Lenguas y tradi-ciones orales de la Amazonia. Diversidad en peligro?* Havana: UNESCO/Casa de las Américas. (p. 71-121).
- Widlok, Thomas. (2005) “Cross-disciplinarity in the documentation work of anthropologists and linguists” In Austin, P. K. (ed) *Language Documentation and Description* V.3. London: ELDP/SOAS (p. 12-21)