

# DIGITALISASI TEKS LISAN BAHASA DHAO<sup>1</sup>: Sebuah Metode Dokumentasi dan Revitalisasi Modern

Jermy Imanuel Balukh

STIBA Cakrawala Nusantara Kupang  
Jalan Bunga Jati No. 88 Oesapa Kupang 85228  
Telepon 0380-8553669  
[jbalkh@yahoo.com.au](mailto:jbalkh@yahoo.com.au)

## ABSTRAK

Makalah ini membahas perihal revitalisasi bahasa Dhaio, sebuah bahasa dengan sekitar 3000 penutur di Pulau Ndao, Provinsi NTT. Pembahasan difokuskan pada pemanfaatan dan digitalisasi teks lisan untuk berbagai tujuan. Agar teks lisan dapat dimanfaatkan, maka digunakan teknologi digitalisasi – perekaman, anotasi, penyuntingan, dan lain-lain. Hasilnya, konstruksi kalimat, tataurutan kata, intonasi, data leksikal dan berbagai ekspresi diperoleh secara alamiah. Untuk tujuan pendidikan, maka dilakukan penyuntingan rekaman yang nantinya menyertai teks tertulis hasil transkripsi. Dengan digitalisasi, upaya pengembangan ranah dan fungsi bahasa Dhaio dapat tercapai. Transmisi bahasa pun tidak hanya mengandalkan kemampuan para penutur, melainkan pengalaman berbahasa secara alamiah yang terdokumentasi dengan baik.

Kata kunci: kepunahan, dokumentasi, teks, digitalisasi, revitalisasi.

## ABSTRACT

This paper deals with the revitalization of Dhaio, a language spoken by about 3000 people in Ndao Island, East Nusa Tenggara Province. The discussion focuses on the use and digitalization of oral text for multipurpose function. To make the oral text useful, the digitalization technology is used — recording, annotation, editing, etc. As the result, sentence construction, word order, intonation, lexical data, and variety of expressions are naturally produced. For educational purposes, recording editing must be done in order that the edited recording can usefully accompany the written text of transcribed data. By digitalization, the effort of developing domains and functions of Dhaio can be reached. The language transmission depends not only on the speakers' competence, but also the natural speaking experience which is well-documented.

Keywords: endangerment, documentation, text, digitalization, revitalization.

## PENDAHULUAN

Revitalisasi bahasa dimaknai sebagai upaya menciptakan bentuk dan fungsi baru tertentu terhadap suatu bahasa yang terancam punah. Hal ini bertujuan agar penggunaan bahasa tersebut meningkat, bahkan pengguna bahasa pun bertambah. Revitalisasi bahasa meliputi, tidak hanya upaya memperluas sistem linguistik dari suatu bahasa minoritas, tapi juga menciptakan ranah baru dalam penggunaannya oleh tipe penutur yang baru pula karena, menurut banyak ahli, hilangnya ratusan bahkan ribuan bahasa merupakan suatu bencana intelektual (King, 2001:5–9).

Hampir 20 tahun terakhir, revitalisasi bahasa menjadi sebuah fokus studi yang penting di kalangan pakar linguistik. Bidang ini menjadi sangat penting karena, memang, bahasa merupakan sisi yang tidak terpisahkan dari kehidupan masyarakat. Studi ini berkembang dengan pesat dan menyebar luas dalam bingkai dokumentasi karena tujuannya yang mulia, yaitu mengembangkan, menciptakan ranah dan fungsi baru, bahkan menyelamatkan bahasa. Untuk mencapai tujuan dimaksud, maka teknologi memainkan peran yang sangat signifikan karena hasil dokumentasi untuk tujuan revitalisasi harus memiliki daya tahan jangka panjang. Dengan demikian, generasi berikutnya masih dapat menikmati hasilnya, bahkan hingga generasi yang mungkin tidak bisa lagi berbicara dalam bahasa tersebut (lihat Himmalmann, 2006a).

Makalah ini membahas perihal revitalisasi bahasa Dhaio, sebuah bahasa yang dituturkan oleh sekitar 3000 orang di Pulau Ndao, Kabupaten Rote-Ndao, Provinsi NTT. Pembahasan dalam makalah ini difokuskan pada pemanfaatan teks lisan untuk berbagai tujuan. Selain itu, berbagai hal teknis mengenai pemrosesan teks lisan secara digital agar menjadi bahan siap pakai untuk tujuan tertentu pun akan dibahas. Data yang digunakan dalam makalah ini diambil dari proyek dokumentasi bahasa Dhaio<sup>2</sup> yang telah

<sup>1</sup> Makalah ini dipresentasikan pada Seminar Nasional Bahasa Ibu II, tanggal 27-28 Februari 2009 di Pascasarjana Universitas Udayana, Denpasar Bali, dengan judul *Representasi Teks Lisan dalam Revitalisasi Bahasa Dhaio*.

<sup>2</sup> Judul proyek: *Documenting Folk Tales and Procedural Texts in Dhaio dengan dukungan dana dari The Endangered Language Fund (ELF) dan Bill Bright Award 2008*, New Haven, USA.

menghasilkan baik rekaman audio dan maupun rekaman video mengenai cerita rakyat dan teks prosedural (proses mengerjakan sesuatu). Kedua hal tersebut sangat penting karena cerita rakyat dapat mengungkap nilai budaya dan teks prosedural mengungkap kehidupan ekonomi dan sosial masyarakat.

## PEMBAHASAN

### Sekilas mengenai Bahasa Dhao dan Kondisinya

Nama pulau dan bahasa yang dikenal luas dengan istilah *Ndao* sebenarnya mengikuti pelafalan bahasa yang lebih dominan di daerah itu, yaitu bahasa Rote. Secara linguistik, bahasa Dhao tidak memiliki bunyi /nd/. Bunyi yang digunakan adalah bunyi dental bersuara yang agak retrofleks dan afrikat yang dilambangkan dengan [d], menggunakan ortografi /dh/ (Grimes, 2006). Sementara itu, secara historis, istilah *dhao* berasal dari nama sebuah pohon kecil penghasil warna hitam, *dhau*<sup>3</sup> ‘tarum’, dengan mengikuti pelafalan salah satu dialek Hawu di pulau Sabu<sup>4</sup> (Balukh, 2007). Menurut Walker (1982:57), selain bahasa Hawu, bahasa Dhao merupakan satu bahasa di Indonesia timur yang memiliki keunikan, yakni empat fonem implosif yang secara ortografis ditulis dengan /b’/, /d’/, /j’/, dan /g’/. Walaupun secara genetis Dhao dan Hawu serumpun, namun secara tipologis Dhao mendapat pengaruh yang signifikan dari Rote. Hal ini disebabkan oleh faktor geografis, ekonomi, dan politik.

Bahasa yang memiliki penutur yang banyak belum bisa dikatakan bahasa yang hidup (Crystal, 2000). Kondisi suatu bahasa hanya dapat dinilai dengan menggunakan parameter tertentu. UNESCO (2003) menetapkan sembilan faktor yang dapat digunakan untuk menilai kondisi suatu bahasa. Faktor-faktor tersebut adalah (1) transmisi bahasa antargenerasi, (2) jumlah penutur, (3) proporsi penutur dari total penduduk, (4) ranah bahasa yang ada, (5) tanggapan terhadap ranah dan media baru, (6) bahan-bahan untuk pendidikan dan keaksaraan bahasa, (7) sikap dan kebijakan bahasa secara institusional dan pemerintah, termasuk status dan penggunaan resmi, (8) sikap anggota masyarakat, dan (9) jumlah dan kualitas dokumentasi. Kesembilan faktor tersebut dapat digunakan secara bersamaan untuk mengukur vitalitas suatu bahasa dan seberapa jauh pentingnya mendokumentasikan suatu bahasa. Untuk memperoleh gambaran mengenai kondisi bahasa Dhao, maka deskripsi ringkas berdasarkan kesembilan faktor di atas ditampilkan di bawah ini.

Bahasa Dhao sebenarnya memiliki sekitar 7000-an penutur yang tersebar di Provinsi NTT (Grimes, 2006; Balukh, 2007). Akan tetapi, hingga saat ini tidak terdapat data statistik yang pasti mengenai jumlah penutur aslinya. Di pulau Ndao sendiri, jumlah penduduk hanya 2.739 (BPS, 2007). Hal ini memberi indikasi bahwa orang asli Ndao yang bermukim di pulau Ndao hanya berkisar 39,13%, dan selebihnya berada di daerah lain. Kecenderungan berkurangnya penduduk di pulau tersebut lebih banyak dipengaruhi oleh faktor ekonomi (Balukh, 2006; 2007)<sup>5</sup>. Masyarakat Ndao juga multilingual – paling sedikit menguasai dua bahasa, yakni Dhao dan Melayu Kupang<sup>6</sup>. Banyak dari mereka bisa berbicara dalam bahasa Rote karena kehidupan ekonomi mereka tergantung pada Rote. Situasi ini menyebabkan transmisi bahasa antargenerasi berada pada posisi tidak aman (*unsafe*). Memang, anak-anak masih menggunakan bahasa ibu mereka di rumah, di gereja, di luar kelas, namun tidak sedikit orang tua berbicara dengan anak-anak dalam bahasa Melayu Kupang tanpa melihat situasi, tempat, dan waktu. Hal ini dapat dibuktikan dengan sedikitnya orang tua yang bisa bercerita dalam bahasa Dhao. Selain itu, jarang ditemui penatua gereja yang mampu berkhotbah atau memimpin ibadah dengan menggunakan bahasa Dhao. Selain itu, belum dikembangkannya bahan ajar muatan lokal dalam bahasa Dhao membuat penggunaan bahasa ini justru semakin minim tanpa regenerasi yangersistem.

Sejak tahun 2000-an, tim SIL<sup>7</sup> mulai mengembangkan terjemahan Alkitab<sup>8</sup> ke dalam bahasa Dhao dan menghasilkan beberapa kitab. Walaupun penggunaan kitab-kitab tersebut belum maksimal, namun paling tidak ortografi yang digunakan sangat membantu, khususnya dalam hal pengenalan bahasa tulis. Ketiadaan bahan tertulis, seperti bacaan atau teks, kamus<sup>9</sup>, dan tata bahasa yang memadai menyebabkan kurangnya perhatian masyarakat terhadap bahasa tulis, walaupun pada dasarnya masyarakat sangat mendukung upaya pemertahanan dan pengembangan bahasa mereka. Di sisi lain, kebijakan pemerintah mengenai upaya revitalisasi bahasa masih pasif – terbatas pada kelengkapan administrasi. Kondisi bahasa Dhao dapat dirangkum dalam Tabel 2.1 di bawah ini.

<sup>3</sup> Daun dari pohon ini digunakan sebagai penghasil warna hitam untuk tenunan, namun sekarang tidak digunakan lagi karena sudah tersedia pewarna modern di pasaran.

<sup>4</sup> Secara genetis, Dhao dan Hawu merupakan satu rumpun bahasa (Grimes, 2006)

<sup>5</sup> Banyak orang pergi atau bahkan pindah ke daerah lain untuk berdagang hasil kerajinan, seperti tenunan dan perhiasan perak.

<sup>6</sup> Banyak orang Ndao juga bisa berbicara dalam bahasa Rote, Hawu (Sabu), dan bahasa Indonesia.

<sup>7</sup> Summer Institute of Linguistics

<sup>8</sup> Kitab suci orang Kristen

<sup>9</sup> SIL sedang mengolah sekitar 1600 kata untuk kamus (lihat [www.sil.org](http://www.sil.org)), juga salah satu tujuan dokumentasi (Balukh, 2008).

Tabel 1. Estimasi Tingkat Ancaman dan Pentingnya Dokumentasi Bahasa Dhao

No	Faktor	Nilai	Keterangan
1	Transmisi bahasa antargenerasi	4 (tidak aman)	Bahasa Dhao digunakan oleh anak-anak hanya pada ranah tertentu, seperti di rumah, di luar kelas, dan di gereja namun tidak maksimal.
2	Jumlah penutur secara absolut	± 3000 orang (39,13%)	Sebenarnya terdapat ±7000-an orang Ndao yang tersebar di NTT, tapi yang bermukim di Ndao tidak mencapai 3000 orang, bahkan berkurang setiap tahun.
3	Proporsi penutur dari total penduduk	4 (mulai terancam)	Hampir semua bisa berbicara dalam bahasa Dhao, tapi hanya sedikit yang bisa bercerita, berkhotbah.
4	Kecenderungan ranah- ranah bahasa yang ada	3 (kurang)	Bahasa Dhao digunakan dalam ranah rumah tangga dan untuk banyak fungsi, tapi bahasa yang dominan (Melayu Kupang) sudah menembus ranah rumah tangga.
5	Tanggapan terhadap ranah dan media baru	1 (minimal)	Bahasa Dhao hanya digunakan dalam upacara resmi tertentu, tapi masih minim dan belum ada media.
6	Bahan-bahan untuk pendidikan dan keaksaraan bahasa	2 (kurang)	Beberapa bahan tertulis, seperti bacaan Injil tersedia, tapi hanya bermanfaat bagi kalangan tertentu. Pendidikan bahasa Dhao belum dimasukkan dalam kurikulum.
7	Sikap dan kebijakan institusional dan pemerintah, termasuk status dan penggunaan bahasa secara resmi	3 (pasif)	Tidak ada kebijakan yang eksplisit. Dalam ranah publik, bahasa Indonesia digunakan.
8	Sikap masyarakat terhadap bahasa sendiri	4 (mendukung)	Kebanyakan penutur bahasa Dhao mendukung pemertahanan bahasa.
9	Jumlah dan kualitas dokumentasi	2 (tidak lengkap)	Terdapat gambaran tatabahasa, daftar kata, dan teks dengan jangkauan terbatas. Terdapat rekaman yang bervariasi.

### Peran Teks Lisan

Pentingnya teks dalam dokumentasi bahasa telah dikenal luas, namun teks hanya mewakili sebagian kecil (mungkin sepersepuluh persen) dari bahan-bahan dokumentasi bahasa yang ada (Himmelmann, 1998 dalam Holton, 2005). Walaupun demikian, rekaman dan transkripsi merupakan hal yang sangat penting dalam dokumentasi. Jika demikian, peran teks lisan, seperti cerita rakyat dan teks prosedural, sangat penting dalam upaya pemertahanan dan revitalisasi bahasa (lihat Holton, 2005). Terdapat dua hal penting yang berkaitan dengan kontribusi teks lisan dalam revitalisasi bahasa. Pertama, teks lisan dapat memberikan gambaran bagi model tertulis yang nantinya bermanfaat dalam pengembangan bahan tertulis. Kedua, dengan adanya rekaman, teks lisan dapat didengar sebagai cerminan pengalaman asli para penutur (lihat Holton, 2005). Dengan adanya teks lisan yang terrekam dengan baik, transmisi bahasa, khususnya cerita rakyat, tidak hanya bergantung pada kemampuan ‘menuturkan’, tapi rekaman menjadi wadah baru dalam penggunaan bahasa.

Dalam hal revitalisasi bahasa, perekaman audio, video, dan transkripsi teks lisan sangat penting. Audio dan video merupakan alat yang umumnya digunakan dalam mendokumentasikan cerita-cerita lisan. Akan tetapi, hal tersebut bukan alat utama dalam penyebaran teks. Teks lisan akan lebih bermanfaat apabila disebar dalam bentuk tulisan, seperti transkripsi dan rekaman yang telah disunting.

Dalam hal pengembangan bahan tertulis, teks lisan memainkan peran yang sangat penting. Teks lisan dapat menghasilkan pengucapan yang alamiah dan baik sehingga digunakan dalam pengembangan ortografi bahasa tersebut (Seifart, 2006). Teks lisan juga menghasilkan konstruksi kalimat yang alamiah, walaupun

secara linguistik belum tentu beraturan. Kalimat-kalimat hasil transkripsi menjadi basisdata penting dalam pengembangan tatabahasa (Mosel, 2006). Berkaitan dengan data leksikal, teks lisan sangat potensial dalam menghasilkan banyak kata dengan berbagai variasi, diksi, dan bentuk dalam waktu singkat. Selain itu, kata-kata yang dihasilkan menggambarkan berbagai makna dan fungsi dalam bahasa tersebut, dan ini sangat penting dalam membangun basisdata leksikal untuk kamus (Haviland, 2006).

### **Peran Teknologi: Format dan Aksesibilitas**

Teknologi sendiri memang tidak dapat menyelamatkan bahasa yang terancam punah, tapi banyak ahli mengakui bahwa teknologi menjadi alat yang secara tidak langsung memiliki kontribusi yang besar terhadap upaya revitalisasi bahasa (Holton, 2005). Agar teks lisan dapat dimanfaatkan untuk berbagai tujuan, termasuk membuat sistem tulis suatu bahasa, maka peran teknologi sangat penting. Sebagaimana tujuan dokumentasi adalah untuk menyediakan bahan-bahan yang dapat bertahan dalam jangka panjang, maka persiapan, proses, hingga hasil dokumentasi harus berkualitas tinggi. Misalnya, rekaman harus masih bisa didengar dengan jelas, tulisan masih bisa dibaca atau dibuka oleh program komputer, dan video masih bisa ditonton pada seratus tahun mendatang, bahkan hingga generasi berikutnya yang mungkin tidak bisa lagi berbicara dalam bahasa tersebut (Himmelman, 2006a).

### **Rekaman**

Hingga sekarang, bahasa tulis belum berkembang di Ndao, walaupun sudah terdapat bahan tertulis, seperti terjemahan Alkitab yang secara tidak langsung memberi gambaran mengenai ortografi dan tatabahasa. Karena itu, yang menjadi andalan utama dalam memperoleh data bahasa untuk kepentingan penelitian adalah tuturan lisan<sup>10</sup>. Tuturan lisan bisa dalam bentuk cerita rakyat, penggunaan alat, proses kerja, percakapan sehari-hari, dan lain-lain. Agar tuturan lisan dapat dijadikan data, maka metode yang lebih canggih dan efektif adalah perekaman, baik audio maupun video. Perekaman memang merupakan metode yang paling universal dalam dokumentasi bahasa (Holton, 2005).

Perekaman yang dilakukan dalam proyek dokumentasi bahasa Dhao oleh Balukh (2008) difokuskan pada cerita rakyat dan teks prosedural, dengan tujuan membuat kumpulan teks dan kamus pengantar. Dua fokus rekaman di atas tentu belum lengkap, namun dapat menjadi dasar untuk penelitian lanjutan. Untuk merekam data yang dibutuhkan, maka beberapa alat digunakan, antara lain perekam (*walkman*), kaset, mikrofon eksternal, kabel dobel-steker, dan alat pendukung lainnya. Pilihan terhadap *walkman* sebagai alat rekam dalam proyek dokumentasi ini memiliki beberapa alasan, antara lain (1) alat tersebut kecil dan mudah dibawa, menggunakan kaset pita yang mudah diperoleh di pasaran, menggunakan baterai, dan mudah dioperasikan, (2) di Ndao tidak terdapat persediaan listrik (PLN)<sup>11</sup>, sehingga tidak bisa menggunakan komputer atau perekam digital, dan (3) perekam digital sulit diperoleh di pasaran, kalau ada, sangat mahal harganya. Walaupun demikian, perekam digital, seperti *ZoomH4* digunakan pada saat perekaman dilakukan di Kupang. Alat rekam apapun yang menjadi pilihan bukan hal yang paling utama. Banyak orang memilih alat tertentu karena memiliki pertimbangan tertentu, seperti kondisi lapangan, ketersediaan alat di pasaran, alokasi dana, atau pertimbangan lain. Hal yang paling penting dari rekaman adalah kualitas suara dan format rekaman yang kuat dan tahan lama (ingat bahwa rekaman akan digunakan untuk berbagai tujuan).

Untuk menghasilkan kualitas rekaman yang baik, maka dipilih perekam yang memiliki fasilitas (1) mikrofon external, (2) menu yang meminimalisasi suara bising, (3) menu penghitung waktu (*counter*), dan (4) *headphone*. Fasilitas-fasilitas tersebut bermanfaat dalam proses perekaman dan digitalisasi<sup>12</sup>. Rekaman bisa menghasilkan suara yang jelas dan baik apabila digunakan mikrofon eksternal yang baik pula<sup>13</sup>, dan bukan mikrofon internal yang biasanya terdapat pada perekam tersebut. Menu untuk meminimalisasi suara bising sangat penting karena biasanya rekaman dengan kaset biasa menghasilkan bunyi getaran. Penghitung waktu pada perekam sangat bermanfaat untuk menentukan posisi rekaman tertentu pada sebuah kaset. *Headphone* berguna untuk mengontrol suara rekaman atau hasil rekaman. Selain itu, menu tersebut bermanfaat untuk menghubungkan perekam dan komputer dengan kabel dobel-steker pada saat proses digitalisasi. Untuk menghasilkan format digital yang baik, maka digunakan perangkat lunak *Audacity* dengan kualitas 16 bit dan 48 kHz. Ini merupakan format untuk kualitas CD, mengingat teknologi DVD belum menjangkau daerah Ndao. Akan tetapi, untuk kepentingan jangka panjang, format standar untuk DVD, yakni 24 bit dan 96 kHz akan dilakukan. Hal ini mesti mendapat perhatian, mengingat di masa yang akan datang,

<sup>10</sup> Ini juga berlaku bagi semua bahasa yang tidak memiliki tradisi tulis.

<sup>11</sup> Beberapa warga memiliki generator listrik, namun hanya digunakan beberapa jam pada malam hari untuk penerangan rumah tangga.

<sup>12</sup> Hal-hal teknis secara mendalam mengenai perekaman tidak dibahas dalam makalah ini.

<sup>13</sup> Model dan tipe mikrofon bisa dipilih sesuai kebutuhan

teknologi CD akan ditinggalkan dan masyarakat akan beralih ke teknologi yang lebih canggih, yakni DVD (lihat Holton, 2005).

Perekam lain yang juga sering digunakan dalam dokumentasi bahasa adalah *mp3 recorder* dan *minidisc*. Akan tetapi, biasanya alat-alat seperti ini tidak direkomendasikan karena walaupun memiliki kualitas suara yang baik, namun format rekamannya kurang kuat, sehingga akan cepat punah atau rusak. Adapun alat rekam lain yang berkualitas tinggi dan biasanya direkomendasikan dalam proyek dokumentasi yang besar dan berskala internasional adalah *Edirol*, *ZoomH4*, dan *flash recorder*. Alat-alat rekam ini memiliki fasilitas tertentu yang dapat menghasilkan kualitas suara yang baik dan format yang kuat. Pada alat-alat tersebut tersedia pilihan format yang dapat disesuaikan dengan kebutuhan dan juga fasilitas penangkapan suara yang dapat disesuaikan dengan kondisi saat rekaman (Juke, 2006). Keuntungan lain adalah tidak dibutuhkan perangkat lunak lain untuk proses digitalisasi karena secara otomatis diproses dalam perekam tersebut. Selain itu, ukuran alat tersebut relatif kecil dan ringan sehingga mudah dibawa. Akan tetapi, alat-alat tersebut susah diperoleh di Indonesia, sehingga tidak jarang para linguist memesan atau membelinya di luar negeri, seperti Australia, Jerman, dan Amerika. Alternatif lain yang juga biasa dilakukan adalah proses perekaman langsung ke komputer dengan menggunakan perangkat lunak tertentu, seperti *Audacity* dan *Cool Edit Pro* disertai mikrofon eksternal. Di sini, kualitas suara tergantung pada kualitas *soundcard* komputer dan mikrofon (bdk Holton, 2005).

### **Anotasi**

Tugas dokumentasi tidak berhenti pada saat rekaman telah dihasilkan. Terutama bahasa yang hanya dipakai oleh sekelompok orang, rekaman-rekaman yang dihasilkan tidak dapat diinterpretasi oleh banyak orang dengan tujuan-tujuan profesional, seperti oleh linguist, antropolog, sejarawan, atau masyarakat umum, bahkan oleh generasi berikutnya. Karena itu, rekaman harus disertai dengan informasi lanjutan dalam format yang dapat diakses oleh pengguna bahasa tersebut. Inilah yang disebut anotasi – berbagai informasi mengenai rekaman, termasuk transkripsi (Berndt, 2006:214). Tiga hal pokok mengenai anotasi yang dibahas di sini adalah transkripsi, terjemahan, dan anotasi gramatikal.

### **Transkripsi**

Istilah transkripsi di sini dimaknai sebagai representasi simbolik dari peristiwa tutur yang didokumentasikan. Transkripsi bisa dilakukan secara ortografis, fonemis, atau fonetis. Dalam hal anotasi, peristiwa tutur yang ditranskripsi harus diperlakukan sama dengan apa yang dilisankan (Berndt, 2006:219–232). Transkripsi harus merepresentasikan sebaik mungkin apa yang dituturkan dalam rekaman. Ini termasuk jeda, salah ucap, dan perbaikan atau pengulangan, bahkan bagian tuturan yang tidak dimengerti, bunyi benda, suara batuk, tarik napas, dan lain-lain pun harus ditandai dalam transkripsi (lihat Du Bois, dkk, 1992).

Tujuan utama dari transkripsi adalah untuk merepresentasikan dalam tulisan seluruh aspek yang terdapat dalam rekaman, baik rekaman audio maupun video, yang nantinya memiliki makna fungsional bagi pengguna bahasa tersebut (Du Bois dkk, 1992:3). Transkripsi memberi gambaran secara umum mengenai suatu bahasa, misalnya fonologi, morfologi, sintaksis, semantik, sosiolinguistik, bahkan sosial-budaya masyarakat penutur bahasa tersebut.

Untuk merealisasikan kualitas dan format transkripsi yang baik, maka digunakan perangkat lunak komputer, seperti ELAN<sup>14</sup>. Di era teknologi modern ini, banyak perangkat lunak lain yang bisa digunakan, seperti *FileMakerPro*<sup>15</sup>, *Transcriber* dan lain sebagainya<sup>16</sup>. Akan tetapi, untuk merepresentasikan berbagai aspek yang berkaitan dengan rekaman, maka banyak ahli bahasa lebih memilih ELAN. Hal ini dikarenakan, ELAN dapat menampilkan audio dan atau video bersama dengan anotasinya, keterkaitan waktu dan segmentasi dalam anotasi sangat tepat, menampilkan anotasi yang tumpang tindih dengan jelas, ada lapisan (*tier*) yang tidak terbatas dan bisa diciptakan sendiri, memainkan rekaman sebagian atau seluruh atau segmen tertentu sesuai pilihan, bisa diekspor dan diimpor dengan perangkat lunak lain, seperti *Toolbox*. Selain itu, teks dapat diekspor dalam bentuk *tab-delimited text files* yang bisa dibuka dengan MS.Word, WordPad atau NotePad. Dengan demikian, rekayasa teks tulis dengan mudah dilakukan.

Hal yang paling penting dalam transkripsi adalah pembagian segmen. Segmentasi yang paling banyak diterapkan adalah unit intonasi, yang dikenal dengan kelompok nada, nafas, frasa intonasional, dan

<sup>14</sup> Singkatan dari *EUDICO Linguistic Annotator*, dikembangkan oleh Birgit Hellwig di Max Planck Institute for Psycholinguistics, Nijmegen, Netherland.

<sup>15</sup> Dikembangkan oleh Brad Taylor di Max Planck Institute for Evolutionary Anthropology, Leipzig, Jerman dengan menciptakan *MasterFM Language Coding Database*.

<sup>16</sup> Lihat [www.sil.org/computing](http://www.sil.org/computing) dan [www.lat-mpi.eu/tools](http://www.lat-mpi.eu/tools) untuk referensi lengkap tentang perangkat lunak yang dapat digunakan dalam kegiatan penelitian bahasa dan budaya.

sejenisnya (Himmelmann, 2006). Selain itu, unit makna juga bisa diterapkan. Hal ini tergantung pada kebutuhan peneliti. Unit intonasi memberi gambaran bahwa bahasa lisan berbeda dengan bahasa tulis. Untuk kepentingan analisis secara fonologis, semantis, atau sociolinguistic, bahkan antropologis maka unit intonasi sangat baik untuk diterapkan. Beberapa ahli menggunakan unit makna dengan tujuan agar lebih mudah mendeteksi hal-hal gramatikal dalam bahasa tersebut. Pada ELAN, segmentasi ditandai dengan referen (\ref). Jika unit intonasi digunakan untuk menandai satu segmen dalam transkripsi maka satu segmen sama dengan satu referen, atau dalam ELAN satu referen sama dengan satu teks (\t). Jika unit makna digunakan, maka satu referen bisa menandai dua atau tiga unit intonasi. Pada ELAN, ini ditandai dengan satu referen sama dengan dua atau tiga teks. Walaupun demikian, unit intonasi menjadi prioritas dalam menentukan segmentasi. Hal yang penting dalam menentukan segmentasi adalah bahwa satu segmen tidak boleh lebih dari satu kalimat.

(1) Segmentasi berdasarkan unit intonasi

```

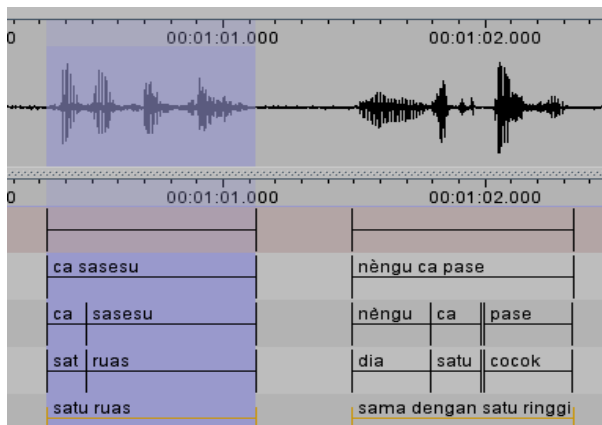
\ref
\t      ca sasesu
\mb     |ca |sasesu |
\fn     satu ruas

\ref
\t      nèngu ca pase
\mb     |nèngu |ca |pase |
\fn     sama dengan satu ringgit

```

Contoh (1) di atas menunjukkan bahwa walaupun secara semantis kedua segmen tersebut merepresentasikan satu makna, namun dituturkan dalam dua kelompok nada, sehingga harus ditranskripsi dengan dua unit intonasi. Unit intonasi biasanya ditandai dengan gelombang suara naik atau turun. Segmentasi di ELAN ditampilkan seperti berikut ini.

Gambar 1. Segmentasi menurut Unit Intonasi pada ELAN



(2) Segmentasi berdasarkan unit makna

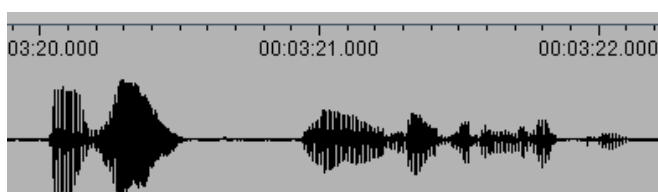
```

\ref
\t      parlu nèngu libu dènge riti
\mb     |parlu | nèngu | libu| dènge| riti|
\fn     mesti dilebur dengan kuningan

```

Contoh (2) di atas menunjukkan satu unit makna dengan sebuah frasa. Sebenarnya dalam rekaman audio terdapat dua unit makna, yakni *parlu* diucapkan terpisah dengan *nèngu libu dènge riti*. Akan tetapi, secara semantis *parlu* menjadi bagian dari unit berikutnya, sehingga ditranskripsi sebagai satu segmen. Segmentasi di ELAN ditampilkan di bawah ini.

Gambar 2. Segmentasi menurut Unit Makna pada ELAN



## Terjemahan

Bahasa Dhao merupakan salah satu bahasa yang dituturkan oleh masyarakat yang terisolasi secara geografis, ekonomi, dan politik. Karena itu, Hasil transkripsi rekaman tidak memiliki manfaat yang lebih luas apabila tidak diterjemahkan ke dalam bahasa yang memiliki penggunaan yang lebih luas, apalagi bahasa Dhao memiliki penggunaan yang terbatas pada kelompok tertentu. Terjemahan yang dikenal dalam sistem dokumentasi bahasa adalah terjemahan bebas (Berndt, 2006:232). Hal yang paling penting dalam terjemahan bebas adalah penentuan bahasa yang digunakan. Dalam konteks ini, terjemahan bebas terhadap hasil transkripsi bahasa Dhao menggunakan bahasa Indonesia dan bahasa Inggris. Pertimbangannya adalah (1) bahasa Indonesia merupakan bahasa resmi dalam negara (Indonesia) dimana bahasa Dhao didokumentasikan atau bahasa nasional yang digunakan dalam pendidikan, dan (2) bahasa Inggris merupakan bahasa resmi dunia yang memungkinkan para ahli dari berbagai penjuru dunia untuk memahami apa yang ditranskripsi. Selain itu, kedua bahasa tersebut juga dipahami secara baik oleh peneliti (Berndt, 2006:234–237). Kemungkinan lain yang dijadikan pertimbangan dalam terjemahan bebas adalah menggunakan bahasa Melayu Kupang yang secara regional lebih dominan digunakan oleh masyarakat di wilayah Timor, termasuk Rote, Ndao, dan Sabu. Akan tetapi, mengingat sistem tulis bahasa tersebut belum berkembang secara luas, sehingga kemungkinan besar tidak memberi kontribusi dalam terjemahan. Sangat penting untuk diperhatikan bahwa bahasa yang dipilih mesti dikuasai oleh penerjemah dan dipahami oleh masyarakat, bahkan generasi berikut dari masyarakat tutur yang mungkin tidak bisa lagi berbicara dalam bahasa tersebut. Selain itu, terjemahan juga perlu menggunakan bahasa yang menjadi target pergeseran bahasa.

Tidak menutup kemungkinan adanya ungkapan idiomatik atau metaforis dalam rekaman. Jika demikian, maka terjemahan bebas sendiri tidak dapat menjelaskan secara detail mengenai makna yang dimaksud dalam transkripsi, sehingga bisa dibantu dengan terjemahan literal. Terjemahan literal bisa dilakukan dengan membuat baris tersendiri dalam anotasi, tapi juga bisa ditulis sebagai tambahan dengan tanda kurung. Selain itu, untuk memberi informasi mengenai makna leksikal daripada kata yang bermakna idiomatik, maka dibuat lapisan khusus untuk pembagian morfem (*morpheme break*). Jika diperlukan, lapisan untuk komentar ditambahkan dalam anotasi.

## Anotasi Gramatikal

Sudah menjadi sebuah standar dalam dokumentasi bahasa untuk memberi informasi gramatikal dengan menyertakan glos morfem. Ini biasanya dilakukan pada saat proses interlinearisasi. Informasi gramatikal bisa mengenai kelas kata, prefiks, klitik, dan lain-lain. Walaupun informasi gramatikal yang disertakan tidak sepenuhnya merupakan gambaran tatabahasa, namun paling tidak menjadi dasar bagi peneliti dalam mengkaji tatabahasa secara mendalam (Berndt, 2006:238–245).

Bahasa Dhao secara morfologis tergolong bahasa isolatif yang hanya memiliki satu afiks (lihat Walker, 1982; Grimes, 2006; Balukh, 2006). Kebanyakan data direpresentasikan dengan satu morfem, satu kata. Karena itu, dalam proses interlinearisasi, terlihat bahwa tidak banyak konstruksi kalimat yang mendapat pemisahan morfem. Walaupun demikian, informasi gramatikal harus ditampilkan dalam anotasi. Biasanya proses interlinearisasi tidak dilakukan pada saat transkripsi di ELAN karena perangkat lunak tersebut tidak memiliki fasilitas untuk itu, kecuali dilakukan secara manual. Interlinearisasi semi-otomatis dilakukan di perangkat lunak lain, seperti *toolbox*. Di bawah ini disajikan salah satu contoh dari bahasa Dhao.

### (3) Interlinear gloss

```
\t      tuku  pakabhe`la
\mb     tuku  pa-kabhe`la
\p      v    pref-adj
\g      tempa CAUS- pelat
\fn     ditempa menjadi pelat
\fe     hit to be flat
```

Contoh (3) di atas menunjukkan bahwa informasi gramatikal ditempatkan pada lapisan kelas kata \p dan glos \g, dimana kata *tuku* adalah verba yang berarti *tempa*, sementara kata *pakabhèla* secara gramatikal dibagi menjadi dua morfem, yakni morfem terikat *pa-* yang merupakan prefiks untuk konstruksi kausatif (CAUS) dan morfem bebas *kabhèla* adalah adjektiva yang berarti *pelat*.

Prefiks *pa-* ‘CAUS’ pada contoh di atas menunjukkan bahwa pemarkah tersebut digunakan untuk menghasilkan makna kausatif yang dilekatkan pada adjektiva. Akan tetapi, belum terdapat informasi mengenai konstruksi kausatif itu sendiri dan adjektiva mana saja yang bisa dilekati *pa-* untuk menghasilkan

makna kausatif. Karena itu, untuk melihat lebih lengkap mengenai persebaran prefiks ataupun informasi gramatikal tertentu maka biasanya dilakukan konkordansi, baik secara terpisah (dengan perangkat lunak tersendiri) maupun secara langsung dalam program *toolbox*.

Proses interlinearisasi, khususnya dalam *toolbox*, tidak bisa terpisahkan dari basisdata leksikal yang berfungsi menyimpan semua data leksikal dan berbagai informasi mengenai data tersebut. Sebelum interlinearisasi dimulai, terlebih dahulu dilakukan pengaturan agar teks dan basisdata leksikal dapat berkomunikasi dengan baik. Keduanya saling membutuhkan karena teks hasil ekspor dari transkripsi menyediakan stok data bagi basisdata leksikal, dan basisdata leksikal memberi berbagai informasi mengenai data itu, seperti kelas kata, makna, penggunaan, dan sebagainya. Pada saat interlinearisasi, data teks akan berkomunikasi dengan basisdata leksikal untuk mengambil informasi yang dibutuhkan. Apabila informasi yang dibutuhkan oleh teks belum atau tidak tersedia dalam basisdata leksikal, maka pada teks akan diberi tanda asterisk tiga (\*\*\*) . Contoh basisdata leksikal ditampilkan seperti di bawah ini.

#### (4) Basisdata dengan model MDF<sup>17</sup>

\lx	<b>pudhi</b>
\ps	adj
\pn	adj
\ge	<i>white ; silver</i>
\gn	putih ; perak
\xv	<i>èdhi mulai libu doi pudhi</i>
\xe	<i>we begin to melt silver</i>
\xn	kita mulai melebur uang perak
\nt	

Basisdata leksikal tidak hanya memberi berbagai informasi mengenai data leksikal untuk kepentingan interlinearisasi, tapi juga digunakan untuk membangun kamus, mulai dari yang paling sederhana, seperti daftar kata atau glosarium, hingga yang lebih kompleks, seperti kamus besar, ensiklopedia, dan lain-lain<sup>18</sup>. Salah satu contoh hasil ekspor dengan menggunakan MDF seperti di bawah ini.

#### (5) Hasil ekspor MDF

<b>pudhi</b>	<i>adj. putih, perak.</i> white, silver.
<b>èdhi mulai libu doi pudhi</b>	
<i>kita mulai melebur uang perak</i>	
<i>we begin to melt silver</i>	

Hasil transkripsi dari rekaman teks lisan tidak hanya menghasilkan teks tertulis untuk bahan bacaan dan sebagai dasar untuk mengkaji tatabahasa, tapi juga bermanfaat untuk membangun kamus yang disertai berbagai informasi, seperti penggunaan, contoh kalimat, padanan terjemahan, dan lain-lain.

### Penyuntingan

Data rekaman yang dihasilkan tidak selalu baik, melainkan selalu terdapat kesalahan ucap, pengulangan, campur kode atau alih kode, dan sebagainya. Karena itu, dibutuhkan penyuntingan agar rekaman dapat digunakan untuk berbagai tujuan. Walaupun demikian, rekaman asli harus tetap disimpan sebagai dokumentasi. Memang, tidak selamanya rekaman asli harus dibandingkan dengan hasil transkripsi yang akan dipublikasikan. Akan tetapi, jika rekaman tersedia maka merupakan suatu pandangan yang cemerlang apabila cerita yang ditranskripsi disertai dengan CD audio (Holton, 2005). Di sini, akan dibahas beberapa hal mengenai pemanfaatan teknologi dalam menyunting hasil rekaman teks lisan.

### Delesi atau Pelesapan

Pada rekaman teks lisan, seperti cerita rakyat, prosedural, dan percakapan, para penutur biasanya bercerita secara alamiah, dan tanpa sadar terdapat kata, kalimat, atau tuturan tertentu diucapkan secara berulang-ulang (mungkin dengan maksud memberi penekanan atau penegasan terhadap hal tertentu). Selain itu, pada rekaman sering terjadi kesalahan ucap (*false start*), *filler*, dan lain-lain. Dalam penyuntingan, tentu pengulangan dan kesalahan yang disebutkan di atas tidak dibutuhkan dalam teks, sehingga harus dihilangkan, baik dari teks transkripsi yang diedit maupun rekaman. Perhatikan contoh di bawah ini.

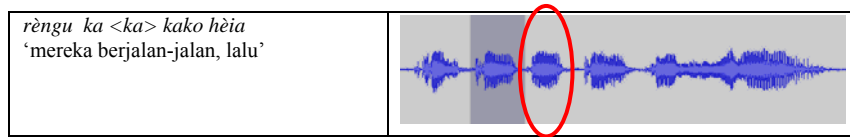
#### (6) Repetisi (rekaman asli)

<sup>17</sup> Multi-Dictionary Formatter

<sup>18</sup> Setelah data diolah dengan MDF, maka bisa dibuka melalui program *Lexique Pro* untuk pemrosesan kamus digital.

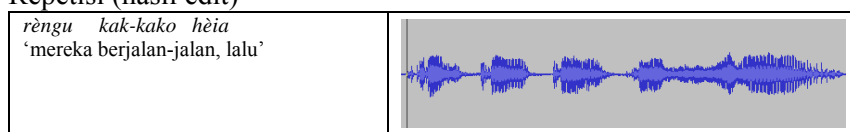


Gambar 3. Repetisi



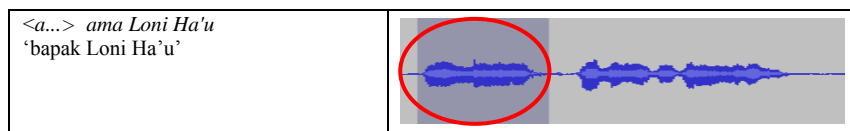
Contoh (6) di atas adalah repetisi yang biasanya diedit atau dihilangkan dalam teks tulis. Transkripsi rekaman asli di atas menunjukkan bahwa penutur bermaksud mereduplikasi kata *kako* 'jalan', namun terjadi pengulangan yang berlebihan. Bagian yang mesti dihilangkan dalam proses penyuntingan rekaman adalah yang bertanda kurung siku, <ka> atau bagian yang dilingkari pada audio di atas. Dengan demikian, teks tersebut menjadi berterima dengan reduplikasi sebagian dari kata *kako* 'jalan', sehingga menjadi *ka-kako* 'jalan-jalan'. Setelah rekaman diedit, pengucapan reduplikasi sebagian tersebut menjadi *kak-kako* yang, menurut penutur asli, merupakan pengucapan yang berterima dan alamiah. Hasil penyuntingan beserta audio ditampilkan dalam contoh (7) berikut ini.

(7) Repetisi (hasil edit)



Hal lain dalam transkripsi yang biasanya dihilangkan adalah *filler*. *Filler* biasanya digunakan pada saat penutur ingin memulai sebuah tuturan atau mengisi jeda, seperti *a...*, *uhm...*, dan lain-lain. Dalam teks tulis, *filler* tidak memberi kontribusi apapun, sehingga harus dihilangkan. Rekaman yang akan dipublikasi bersama teks tulis tersebut tentu tidak memerlukan *filler*. Contoh (8) di bawah ini menggambarkan *filler* di awal tuturan.

(8) Filler (rekaman asli)



*Filler* dalam teks ditandai dengan tanda kurung siku <a...> atau yang dilingkari pada audio, diedit menjadi seperti pada (9) di bawah ini.

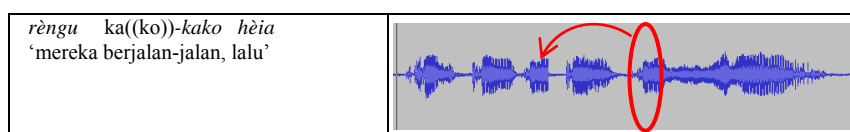
(9) Filler (Hasil edit)



**Penyisipan dan Perluasan**

Para penyunting teks biasanya menyisipkan kata atau bahkan sebagian dari kata tertentu untuk melengkapi bentuk atau konstruksi yang diinginkan. Sementara itu, perluasan dilakukan pada kata atau bentuk tertentu yang dianggap tidak lengkap. Dalam penyuntingan audio, penyisipan belum tentu merupakan perluasan bentuk, tapi perluasan selalu dilakukan dengan mengambil bentuk tertentu dari bagian lain rekaman untuk disisipkan pada bentuk yang sesuai. Perhatikan contoh (10) dan (11) di bawah ini.

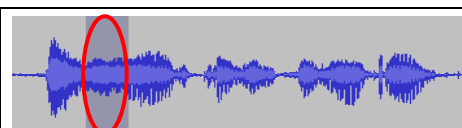
(9) Salinan (diedit dari data (6))



Konstruksi pada (10) di atas merupakan hasil penyuntingan daripada pengulangan pada (6). Reduplikasi yang diharapkan dalam teks sebenarnya adalah *kako-kako* ‘jalan-jalan’. Karena itu, bentuk *ka ka* yang tidak tepat harus dihilangkan dan diisi dengan bentuk *kako*. Sementara bentuk *ka* sudah tersedia, sehingga yang perlu dilakukan adalah bentuk *ko* dalam bagian tersebut disalin dan disisipkan pada *ka*, sehingga menjadi *kako* yang lengkap. Bagian yang disisipkan ditandai dengan dobel kurung pada teks dan anak panah pada audio dalam contoh (10).

Pada contoh (11) di bawah ini dilakukan perluasan terhadap kata *nèngu* ‘dia’ yang ditandai dengan dobel kurung pada teks dan yang dilingkari pada audio.

(10) Perluasan


<p><i>he`ia ((ne`ngu)) nare kahib`i do na ka</i>          ‘lalu dia ambil kambing tadi’</p>	
-------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------

Pada rekaman asli, kata *nèngu* diucapkan *nu*. Secara fonologis, hal ini menjadi analisis yang sangat penting, tapi untuk teks yang akan dipublikasikan, bentuk ini harus direvisi menjadi *nèngu*. Karena itu, perluasan dilakukan dengan cara menyalin bentuk dan bunyi *nèngu* yang sama dari bagian lain dan disisipkan pada posisi *nu*.

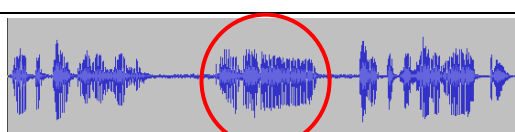
**Pergantian**

Masyarakat Ndao pada umumnya bilingual, bahkan multilingual. Pada tuturan lisan, sering terdapat kata pinjaman, khususnya dari bahasa Indonesia. Hal ini kebanyakan terjadi pada penutur yang lebih muda, sementara penutur berusia tua hanya menggunakan sedikit kata pinjaman. Tampaknya, semakin muda usia penutur, semakin banyak kata pinjaman yang digunakan. Dalam penyuntingan teks, kata pinjaman dari bahasa lain harus diganti. Perhatikan contoh (12) yang direvisi menjadi (13) berikut ini.

(11) Pinjaman (rekaman asli)

<p><i>patao ètu era e`èna na ((atau kah)) èdhi tèke sale</i>  <i>ngaa-ngaa èdhi</i>          ‘berkelahi di tempat itu atau kita salah menyimpan barang-barang kita’</p>	
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------

(12) Pinjaman (hasil edit)

<p><i>patao e`tu era e`èna na ((aado na)) e`dhi te`ke</i>  <i>sale ngaa-ngaa e`dhi</i>          ‘berkelahi di tempat itu atau kita salah menyimpan barang-barang kita’</p>	
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------

Terlihat pada rekaman asli yang ditranskripsi pada (12) di atas bahwa, penutur menggunakan kata *atau* dan partikel *kah* yang merupakan pinjaman dari bahasa Indonesia. Bahasa Dhao sebenarnya memiliki kata dan juga partikel untuk menyatakan makna tersebut, yakni *aado na*. Karena itu, teks dan rekaman harus direvisi. Kata *aado na* dari bagian lain rekaman disalin dan ditempatkan menggantikan kata pinjaman tersebut, seperti yang ditunjukkan oleh (13).

Penyuntingan terhadap rekaman sering mengalami kendala karena hal-hal teknis, seperti tinggi-rendahnya bunyi, tempo, jarak nada, dan lain-lain antara rekaman asli dan hasil saalinan tidak sama. Jika terjadi demikian, maka tugas peneliti adalah menyesuaikan suara dengan menggunakan menu-menu yang tersedia pada perangkat lunak yang digunakan, sehingga hasil penyuntingan tidak terkesan rekayasa tapi benar-benar alamiah.

**SIMPULAN**

Bahasa yang belum memiliki sistem tulis yang berkembang luas seperti Dhao mengandalkan tuturan lisan dalam penyebaran dan transmisi bahasa. Hal ini kurang efektif karena tuturan lisan hanya mengandalkan kemahiran penutur. Akan tetapi, masyarakat semakin bilingual, bahkan multilingual. Karena

itu, tuturan lisan sangat penting didokumentasikan yang nantinya bermanfaat untuk berbagai tujuan dalam upaya revitalisasi bahasa.

Teks lisan yang dihasilkan melalui perekaman memainkan peranan yang sangat penting dalam revitalisasi bahasa. Teks lisan dapat bermanfaat bagi pengembangan sistem tulis dan bahan bacaan, baik untuk kepentingan umum maupun untuk kepentingan pendidikan. Teks lisan menghasilkan data yang bervariasi secara alamiah, sehingga dapat diolah untuk menghasilkan berbagai produk, seperti tatabahasa, kamus, dan lain-lain. Akan tetapi, teks lisan mesti disertai berbagai informasi mengenai apapun yang terdapat dalam rekaman tersebut. Inilah yang disebut anotasi. Anotasi bisa ortografis, fonemis, atau fonetis, tergantung pada tujuan dokumentasi tersebut. Hal yang penting dalam anotasi adalah transkripsi, terjemahan bebas, dan anotasi gramatikal.

Agar rekaman yang dihasilkan dapat digunakan untuk berbagai tujuan, maka teknologi memiliki peran yang sangat penting. Dengan teknologi, proses digitalisasi dapat dilakukan dengan kualitas yang baik dan format yang kuat, sehingga dapat bertahan dalam jangka waktu yang panjang. Teknologi yang dimaksud di antaranya adalah teknologi perekaman, anotasi, dan penyuntingan. Hal yang paling penting adalah bahwa rekaman, anotasi, dan penyuntingan yang dilakukan bertujuan untuk menghasilkan bahan yang dapat dimanfaatkan untuk pengembangan ranah dan fungsi bahasa dalam kerangka revitalisasi bahasa Dhao.

## DAFTAR PUSTAKA

- Balukh, Jermy I. 2006. "Bahasa Ndao dan Budaya Penuturnya yang Terancam Punah" dalam Putra Yadnya, dkk (Ed). *Prosiding Bahasa, Sastra, dan Budaya Austronesia*. Denpasar: Program Pascasarjana Universitas Udayana.
- Balukh, Jermy I. 2007. "Prospek dan Revitalisasi Bahasa Ndao" dalam Mbetje, dkk (Ed). *Makalah Seminar Bahasa Ibu I*. Denpasar: Program Pascasarjana Universitas Udayana.
- Balukh, Jermy I. 2008. Documenting Folk Tales dan Procedural Texts in Ndao. *Proposal Penelitian*. USA: Grant Endangered Language Fund, New Haven.
- Berndt, Eva. S. 2006. Linguistic Annotation dalam Nikolaus P. Himmelmann & Urike Mosel (Ed). *Essentials of Language Documentation*. Berlin, New York: Mouton de Gruyter.
- BPS, 2007. *Rote Ndao dalam Angka*. Kupang: Badan Pusat Statistik Kabupaten Kupang.
- Crystal, David. 2000. *Language Death*. Cambridge: Cambridge University Press.
- Du Bois, John W. dkk. 1992. *Discourse Transcription*. California: Santa Barbara. Department of Linguistics. University of California.
- Grimes, Charles E.. 2006. Hawu and Dhao in Eastern Indonesia; Revisiting Their Relationship. *Makalah Tenth International Conference on Austronesian Linguistics. 17-20 January*. Philippines: Puerto Princessa City, Palawan. <http://www.sil.org/asia/philippines/ical/papers.html>
- Haviland, John B. 2006. Documenting Lexical Knowledge dalam Nikolaus P. Himmelmann & Urike Mosel (Ed). *Essentials of Language Documentation*. Berlin, New York: Mouton de Gruyter Berlin
- Himmelmann, Nikolaus P. 2006a. Language Documentation dalam Nikolaus P. Himmelmann & Urike Mosel (Ed). *Essentials of Language Documentation*. Berlin, New York: Mouton de Gruyter.
- Himmelmann, Nikolaus P. 2006b. Linguistic Annotation dalam Nikolaus P. Himmelmann & Urike Mosel (Ed). *Essentials of Language Documentation*. Berlin, New York: Mouton de Gruyter.
- Holton, Gary. 2005a. *Building The Dena'Ina Language Archive*. Alaska: Alaska Native Language Center
- Holton, Gary. 2005b. *The Representation of Oral Literature and its Role in Language Revitalization*. Alaska: University of Alaska Fairbanks.
- Juke, Anthony. 2006. Rekaman. Materi 1st Training Workshop on Language Documentation Ubud , Bali, 9 – 20 July.
- King, Kendall A. 2001. Language Revitalization Processes and Prospects. Multilingual Matters Ltd.
- Seifart, Frank. 2006. Orthography Development dalam Nikolaus P. Himmelmann & Urike Mosel (Ed). *Essentials of Language Documentation*. Berlin, New York: Mouton de Gruyter.
- Speas, Margaret. NN. Someone Else's Language, On the Role of Linguists in Language Revitalization. Amherst: University of Massachusetts.
- UNESCO. 2003. Language Vitality and Endangerment. Dokumen untuk *International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages di Paris, 10 – 12 Maret*.
- Walker, Alan T. 1982. *Grammar of Sawu*. Jakarta: Badan Penyelenggara Seri NUSA Universitas Atma Jaya