

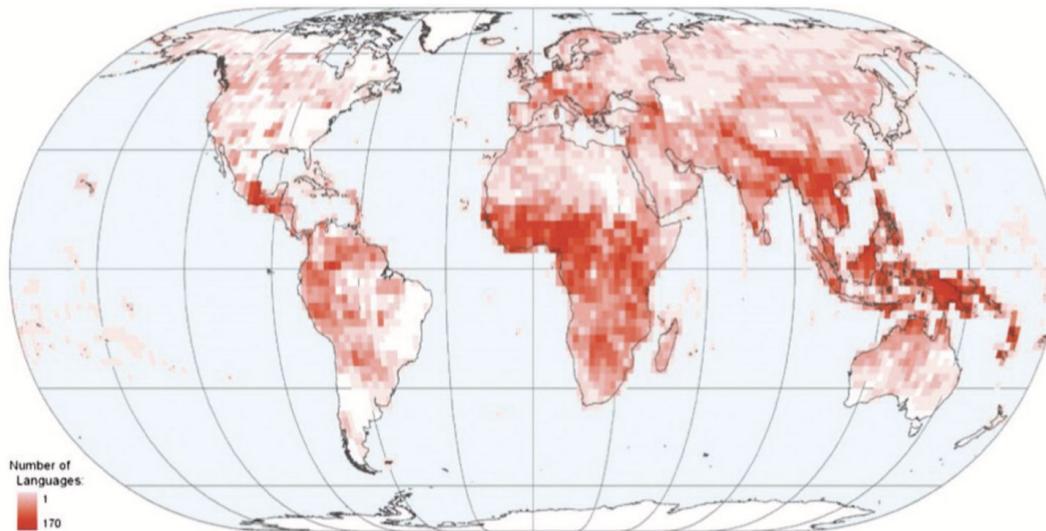
Sonics of Languages

Acoustic Diversity of the Human Languages

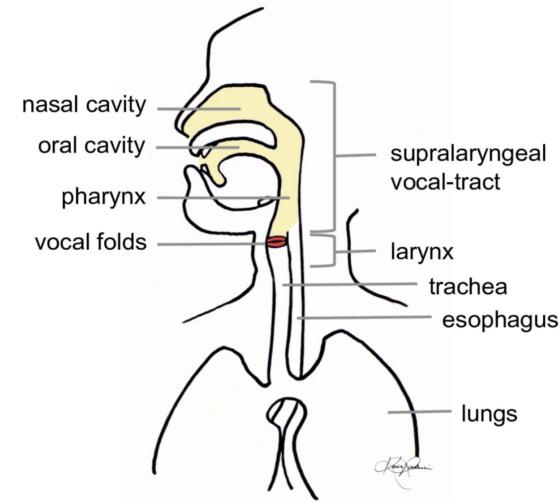
Lydia Krifka-Dobes

Institute of Sonology

Language *Diversity* - and Unity



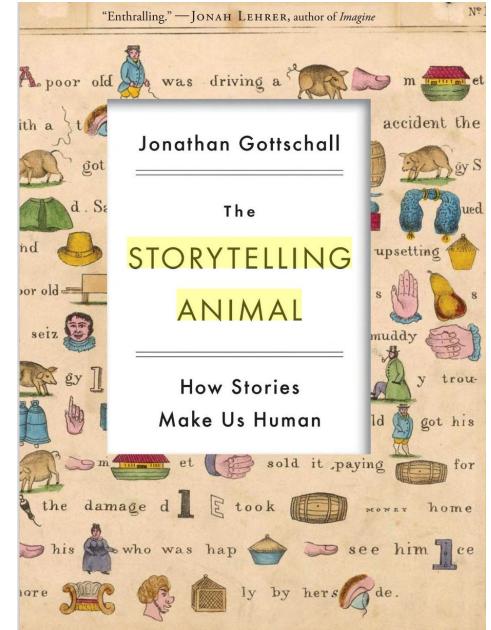
7000 languages
(spoken or signed)
Each using its unique set of phonemes



Biologically similar
production

A unique window into
cultural diversity!

Cultural Diversity – and Unity



Gottschall, Jonathan. 2012. *The storytelling animal. How stories make us human*. New York: Houghton-Mifflin.

"Human life is so bound up in stories that we are thoroughly desensitized to their weird and witchy power."

My story: Visit to Vanuatu



The sounds of languages

Rotokas, Papua New Guinea: 6 consonants, 5 vowels

Central Rotokas

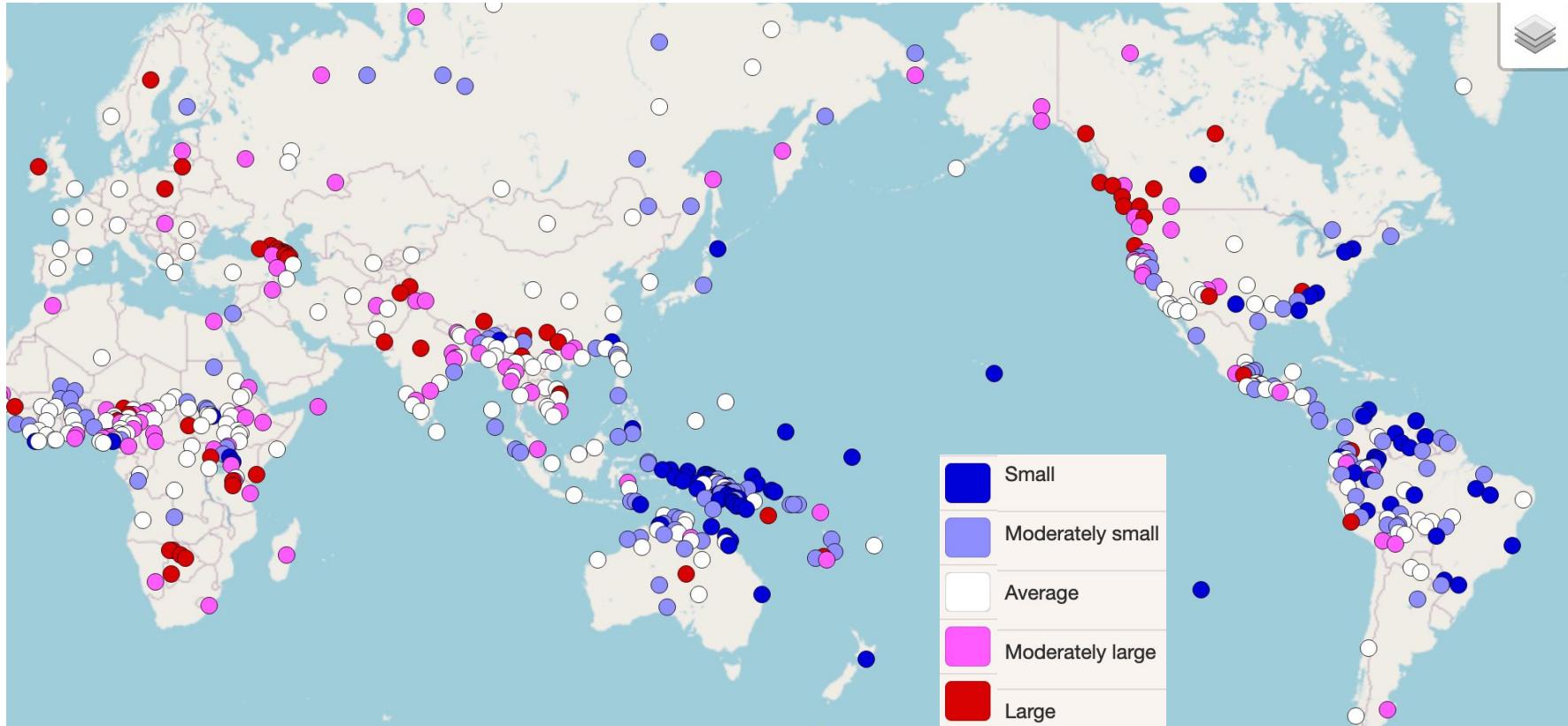
	Bilabial	Alveolar	Velar
Voiceless	p	t	k
Voiced	b ~ β	d ~ r	g ~ γ

	Front	Central	Back
Close	i (iː)		u (uː)
Close-mid	e (eː)		o (oː)
Open		a (aː)	

A very large consonant system (with a tiny vowel system): Abkhaz

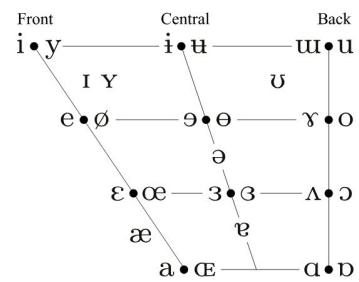
Consonant inventories in the languages of the world

World Atlas of Language Structure WALS



Phonetic features used in languages

VOWELS



DIACRITICS

○ Voiceless	n̥ d̥	.. Breathy voiced	b̥ ḁ	Dental	t̥ d̥
~ Voiced	s̥ t̥	~ Creaky voiced	b̥ ḁ	Apical	t̥ d̥
h Aspirated	tʰ dʰ	~ Lingualobial	t̥̩ d̥̩	Laminal	t̥ d̥
› More rounded	ɔ̥	ʷ Labialized	tʷ dʷ	~ Nasalized	ẽ
⟨ Less rounded	ɔ̥	j Palatalized	tj̥ dj̥	n Nasal release	d̥n
+ Advanced	u̥	v Velarized	tv̥ dv̥	˥ Lateral release	d̥l
- Retracted	e̥	f̥ Pharyngealized	t̥f̥ d̥f̥	˥ No audible release	d̥r
.. Centralized	ɛ̥	~ Velarized or pharyngealized	ɿ̥		
✗ Mid-centralized	ɛ̥	↑ Raised	ɛ̥ (j̥ = voiced alveolar fricative)		
Syllabic	n̥	↓ Lowered	ɛ̥ (β̥ = voiced bilabial approximant)		
~ Non-syllabic	e̥	→ Advanced Tongue Root	ɛ̥		
~ Rhoticity	ɔ̥ ɔ̥	↓ Retracted Tongue Root	ɛ̥		

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2020)

CONSONANTS (PULMONIC)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		t̥ d̥	c j	k g	q ɣ		?
Nasal	m	m̥		n		n̥	n̥̩	ŋ	N		
Trill	B			r					R		
Tap or Flap		v̥		f̥		ɾ̥					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ɟ	xɣ	χʁ	ħ ŋ	hɦ
Lateral fricative				ɬ ɭ							
Approximant		v̥		ɹ̥		ɻ̥	ɺ̥	j	w̥		
Lateral approximant				l̥		ɺ̥	ɺ̥	ɻ̥	L̥		

CONSONANTS (NON-PULMONIC)

Clicks	Voiced implosives	Ejectives
○ Bilabial	b̥ Bilabial	' Examples:
Dental	d̥ Dental/alveolar	p' Bilabial
! (Post)alveolar	f̥ Palatal	t' Dental/alveolar
± Palatoalveolar	g̥ Velar	k' Velar
Alveolar lateral	g̥ Uvular	s' Alveolar fricative

TONES AND WORD ACCENTS

LEVEL	CONTOUR
é or ē	/ Rising
é	\ Falling
é	˥ High
é	˧ Mid
é	˨ Low
é	˩ Extra low
↓	˨˩ Downstep
↑	˥˥ Upstep
	↗ Global rise
	↘ Global fall

The IPA in Music: Ligeti, Aventures

A musical score page from Ligeti's *Aventures*. The score is divided into three staves by a thick horizontal line. The top staff is for the piano (Pf), indicated by a piano icon and dynamic markings like ff and p . The middle staff is for the first voice, labeled 'B' and 'Soprano'. The bottom staff is for the second voice, labeled 'Pf' and '(Hs)'. The vocal parts feature complex IPA transcription above the notes, such as 'dʒɪʃnəns tʃvɛstə' for the soprano and 'dʒɪʃnəns tʃvɛstə' for the bass. The piano part includes vertical bar lines and rests.

Points of Interest: Composing with language

1. Daniel Heller-Roazen, *Echolalias: On the Forgetting of Language*
[Structural metaphors of western thought recited in the Vedic style on Vimeo](#)
2. Lena Herzog, *Last Whispers*, British Museum



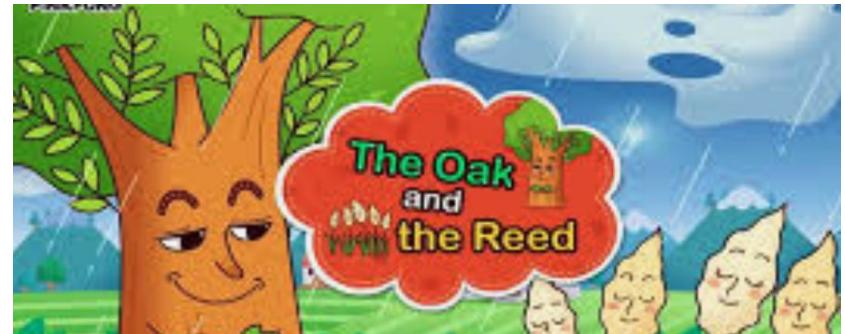
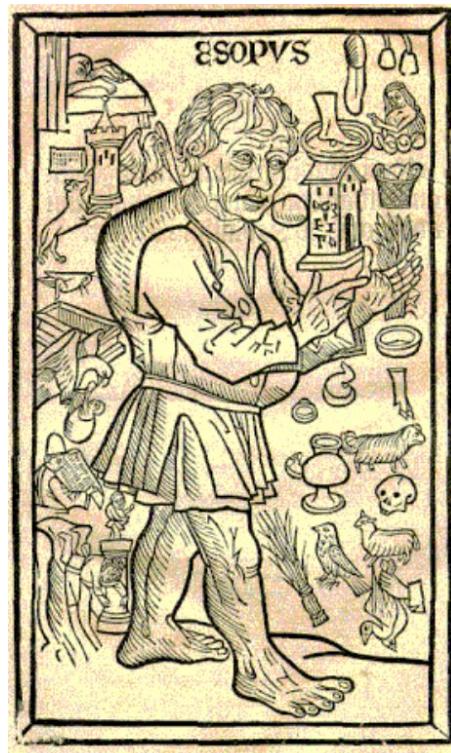
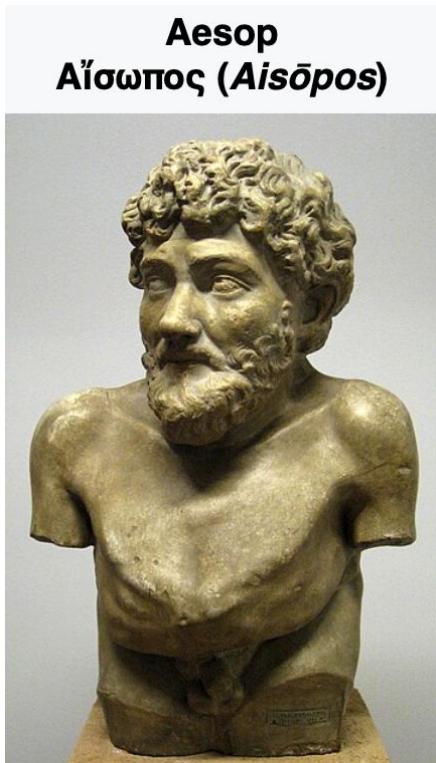
The structure of stories

Vladimir Propp, Claude Lévi-Strauss:
Universal set of story types in myths and fairy-tales.

William Labov and Joshua Waletzki:
Universal structure in oral narratives, containing

- Abstract (title)
- Orientation
- Complicating Action
- Resolution
- Evaluation

The text: Aesop fable “The oak and the reed”



The Oak and the Reed

Very old: Sumerian and Akkadian precursors

Very wide-spread: Chinese proverbs,
Taoist interpretation in Tao te king

Re-told by La Fontaine, *Le chêne et le roseau*

Many modern adaptations as poems,
in art and in music

Translations in many dialects,
minority languages, slang versions



Henri Coutheillas's sculpture of the fable

The Text: The Tree and the Reed

A giant tree stood next to a tiny slender reed near a river.

The tree said to the reed: "Ha, ha, little reed!

Look at **me**. I am strong. No-one can overthrow me. I am the ruler of this land!

And look at **you**. You are just a small feeble reed, going back and forth with the wind".

The reed bowed but said nothing.

Some time later, a huge storm came up.

The giant tree withstood at first and resisted, but then the storm became stronger and stronger.

And finally the storm uprooted the mighty tree, and it fell down with a great thunderous noise.

The reed also felt the mighty storm, but it bent with the wind, this way and that way.

And when the storm was gone, the reed stood upright again.

The reed looked at the fallen tree, and said:

*It is better to yield
when it would be unwise to resist
than to resist stubbornly and be destroyed.*

The Text: The Tree and the Reed

Narration

A giant tree stood next to a tiny slender reed near a river.

The tree said to the reed: "Ha, ha, little reed!

Look at me. I am strong. No-one can
be the ruler of this land!"

Imitated
laughing

Direct
speech

The reed bowed but said nothing.

Some time later, a huge storm came up.

The giant tree withheld at first and resisted, but then the storm became stronger and stronger.

Opposition

Onomatopoeia

And finally the storm uprooted the mighty tree, and it fell down with a great thunderous noise.

The reed also felt the mighty storm, but it bent and, this way and that way.

Moral

And when the storm was gone, the reed straight again.

The reed looked at the fallen tree, and said:

*It is better to yield
when it would be unwise to resist
than to resist stubbornly and be destroyed.*

The Text: The Tree and the Reed

Narration

A giant tree stood next to a tiny slender reed near a river.

The tree said to the reed: "Ha, ha, little reed!

Look at me. I am strong. No-one can
be the ruler of this land!"

Imitated
laughing

Direct
speech

The reed bowed but said nothing.

Some time later, a huge storm came up.

The giant tree withheld at first and resisted, but then the storm became stronger and stronger.

Opposition

Onomatopoeia

And finally the storm uprooted the mighty tree, and it fell down with a great thunderous noise.

The reed also felt the mighty storm, but it bent and swayed, this way and that way.

Moral

And when the storm was gone, the reed stood upright again.

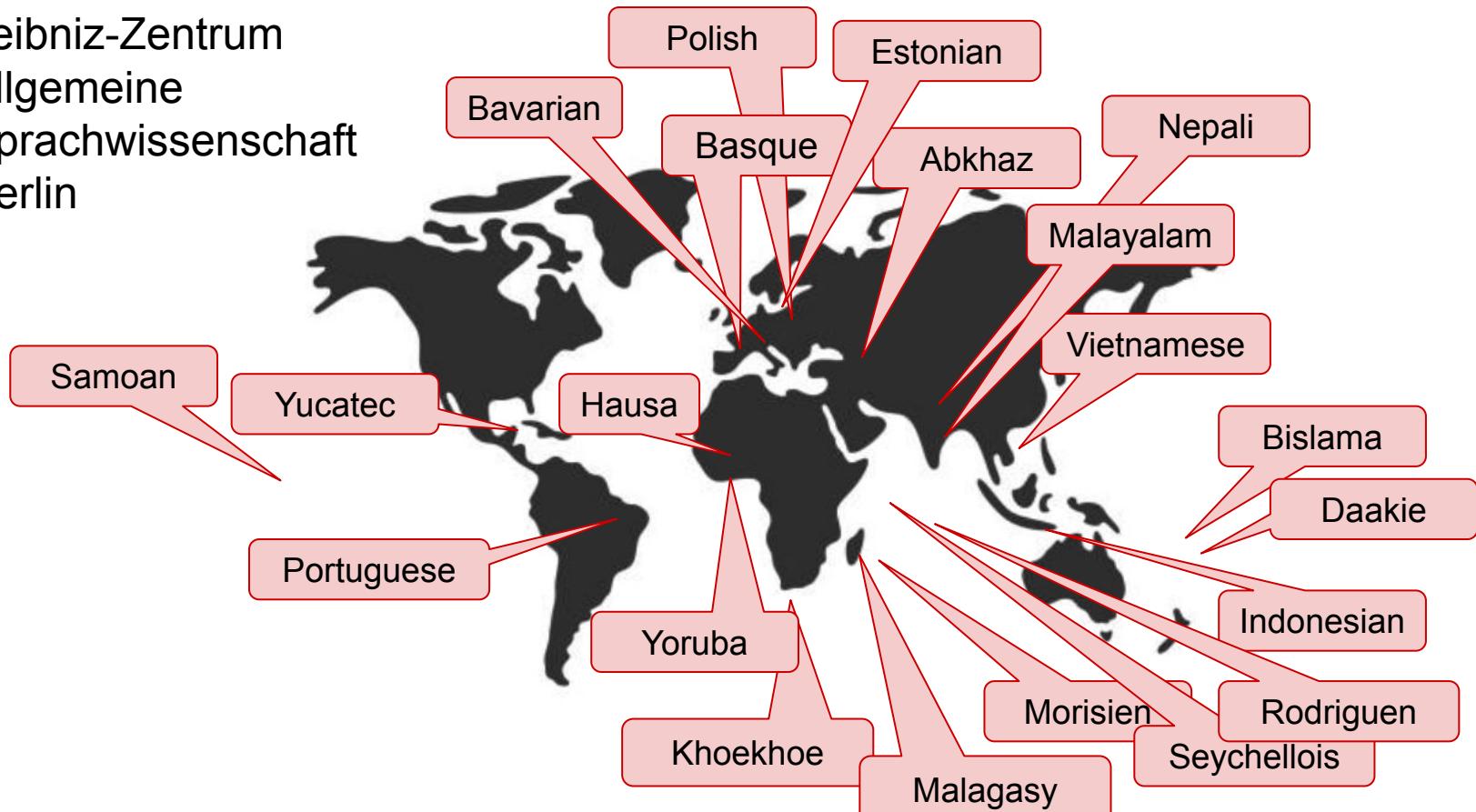
The reed looked at the fallen tree, and said:

It is better to yield

Message: The small and weak can outlive and triumph over the great and strong.

The languages of the recordings

Leibniz-Zentrum
Allgemeine
Sprachwissenschaft
Berlin



Sonics of Languages: Examples

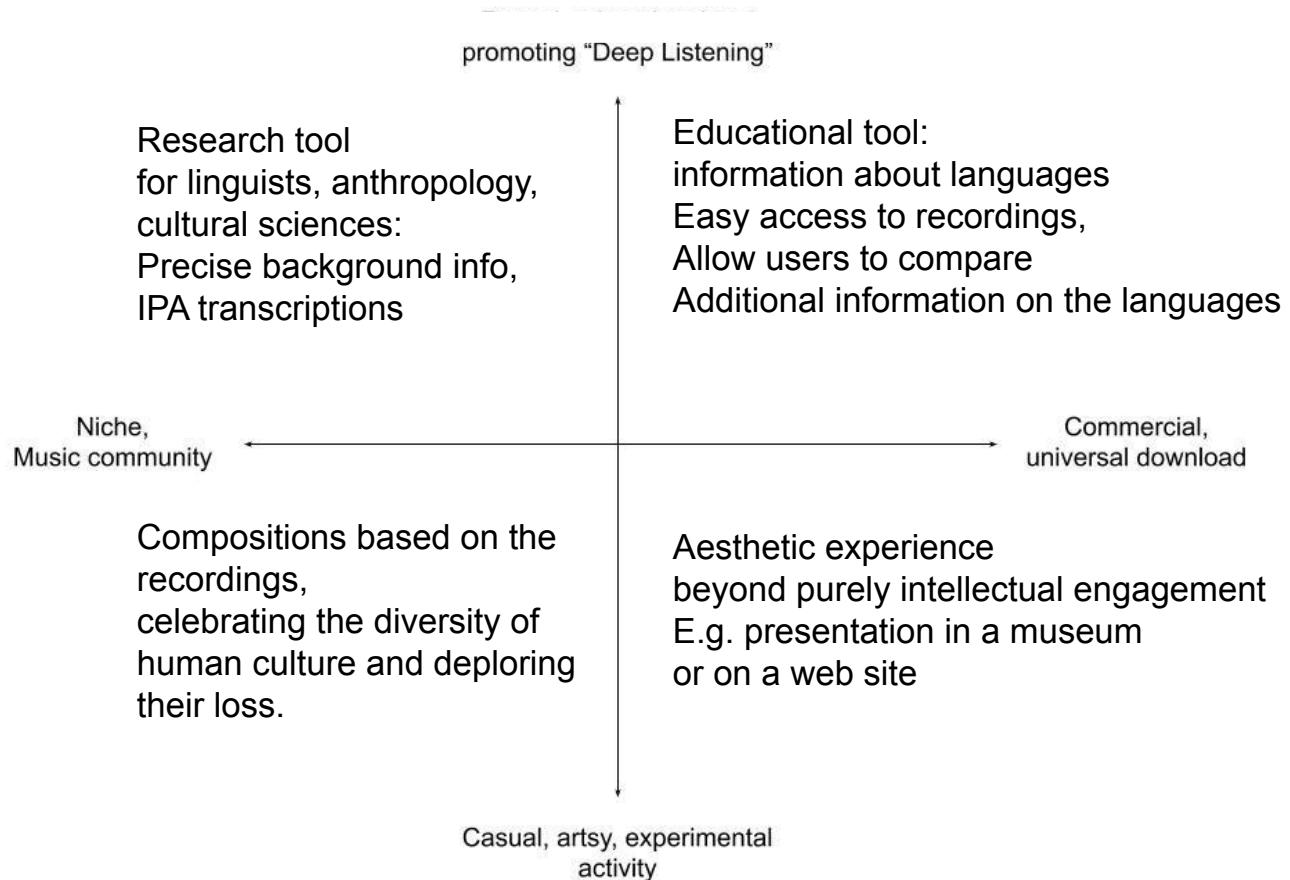
Khoekhoe: clicks



Yucatec: ejectives



Artistic approaches



Aesthetic approaches



Want to do

Compositions based on the recordings,
celebrating the diversity of
human culture and deplored
their loss.

Research tool
for linguists, anthropology,
cultural sciences:
Precise background info,
IPA transcriptions

done

Educational tool:
information about languages
easy access to recordings,
Allow users to compare
Additional information on the languages

Aesthetic experience
beyond purely intellectual engagement
E.g. presentation in a museum
or on a web site

Conceivable
but ambitious

Educational tool

Presentation of languages and sentences on an app,
Including additional information about languages

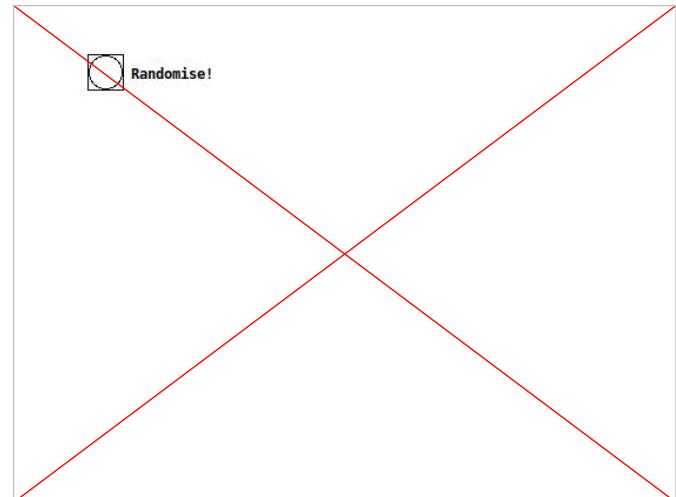
Play

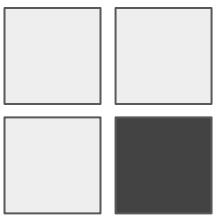
Go to Beginning

Play single Sentence

Repeat Sentence

Repeat Sentence with different Language



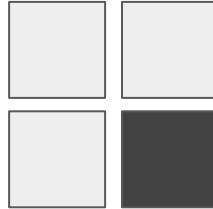


Presentation in a museum

Example: Hörraum in the Humboldt Forum,
Ethnographic Museum of Berlin



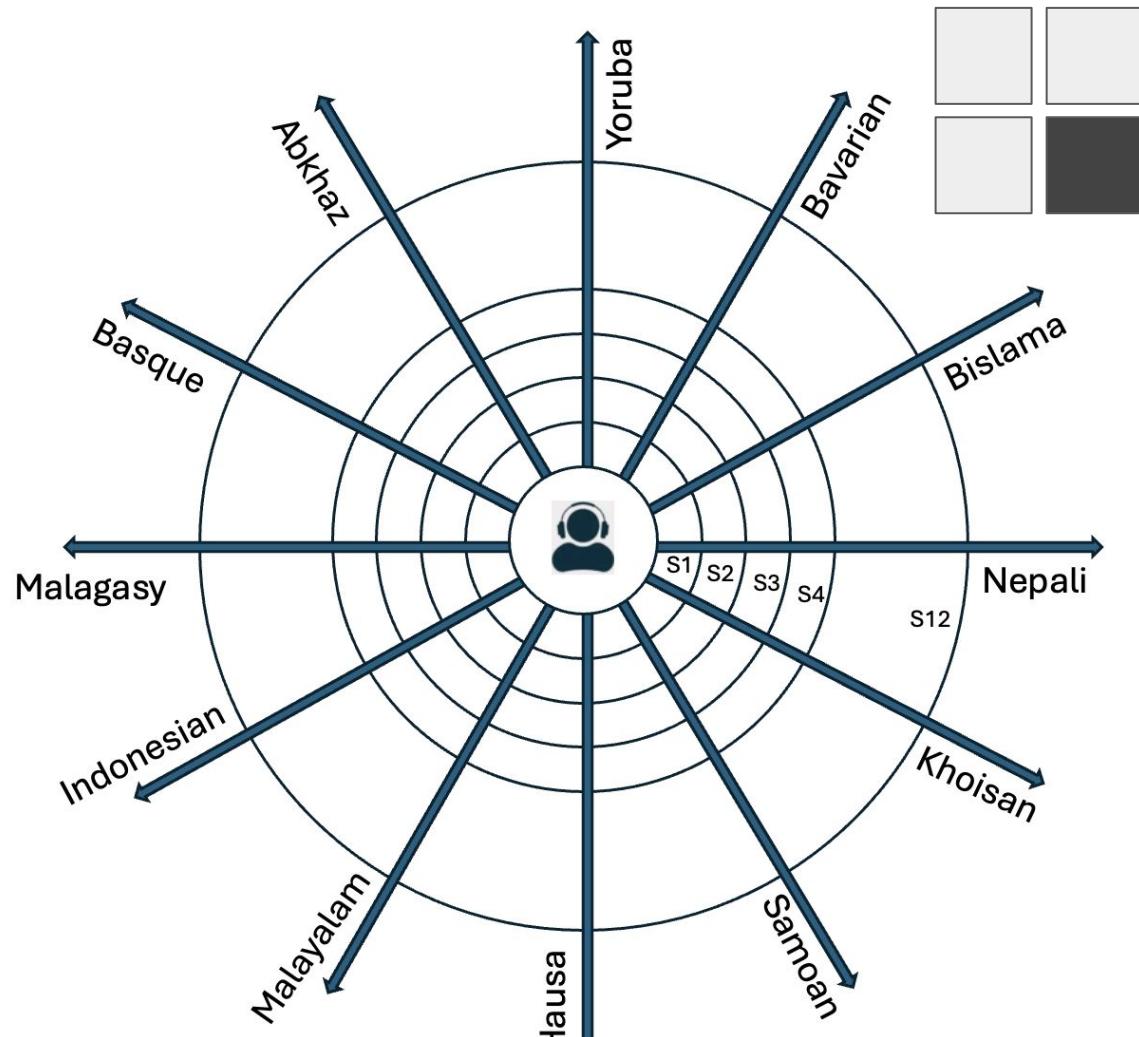
Aesthetic experience: Installation



Walking through a space with sensors (continuation of “Walking the piano”)

Ambisonic earphones (Future plan)

Looking in different directions,
switching from one language
to another one,
sentence by sentence





Abkhazian - Tree_A M S ROUTE FX

Basque - Tree_A M S ROUTE FX

Bavarian - Tree_A M S ROUTE FX

Bislama - Tree_A M S ROUTE FX

Daakie - Tree_A M S ROUTE FX

Estonian - Tree_A M S ROUTE FX

Hausa - Tree_A M S ROUTE FX

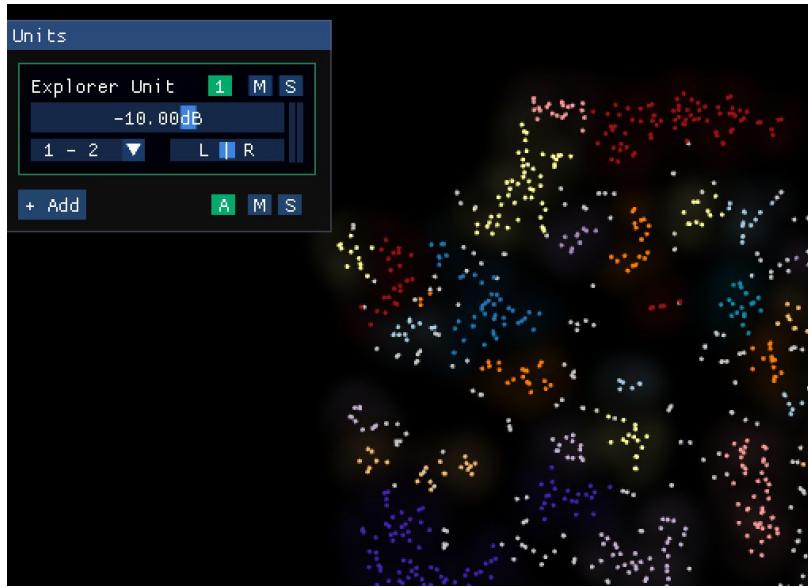
Indonesian - Tree_A M S ROUTE FX

Khoekhoe - Tree_A M S ROUTE FX

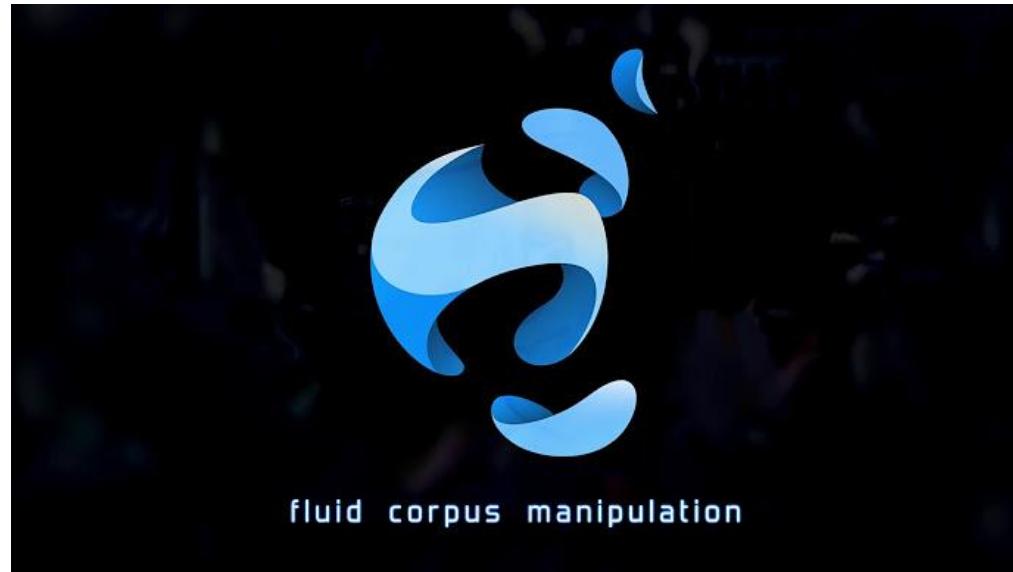
Malagasy - Tree_A M S ROUTE FX

Seychellois - Tree_A [input] Mic/Line 1





Audiomaps with Audiostellar





AudioStellar

Window



AudioStellar v1.4 // Tree and Reed All Sentences.stft.4.json (253 sounds)

File View Settings Tools Help

Link

120

BPM

Units

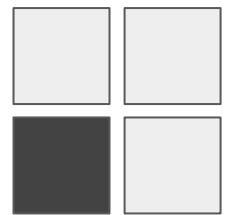
Explorer Unit 1 M S

-10.00dB

1 - 2 ▾ L R

+ Add

A M S

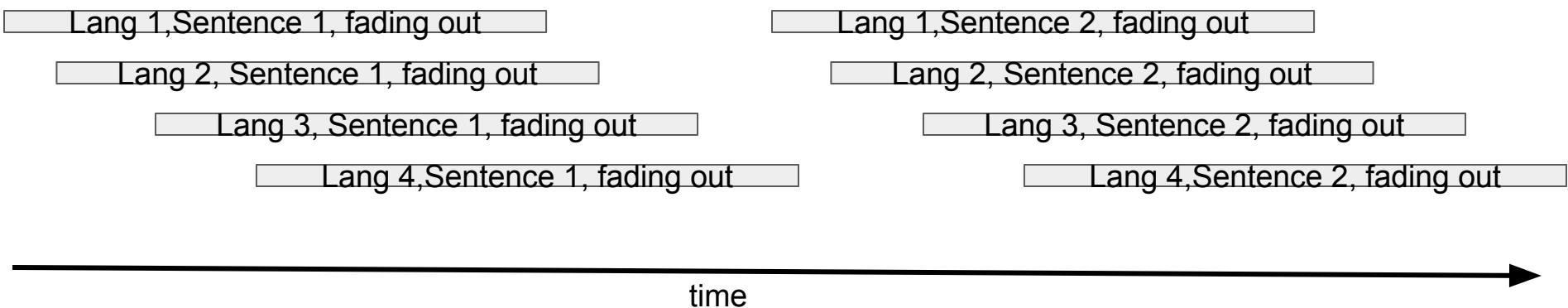


Composition

Speakers of different languages tell the story together,
their individual contributions are weaved together, reflecting the structure of the story

Underlying form of a fugue,
where the theme is the meaning, the voices are the languages

Example: First two sentences, only 4 languages



A giant tree stood next to a tiny slender reed near a river.

The tree said to the reed: **Ha, ha ha**, little reed!

Look at **me**. I am strong. No-one can overthrow me. I am the ruler of this land!

And look at you. You are just a small feeble reed, going back and forth with the wind”.

The reed bowed and said nothing.

SHORT PAUSE

Some time later, a huge storm came up.

The giant tree withstood at first and resisted, but then the storm became stronger and stronger.

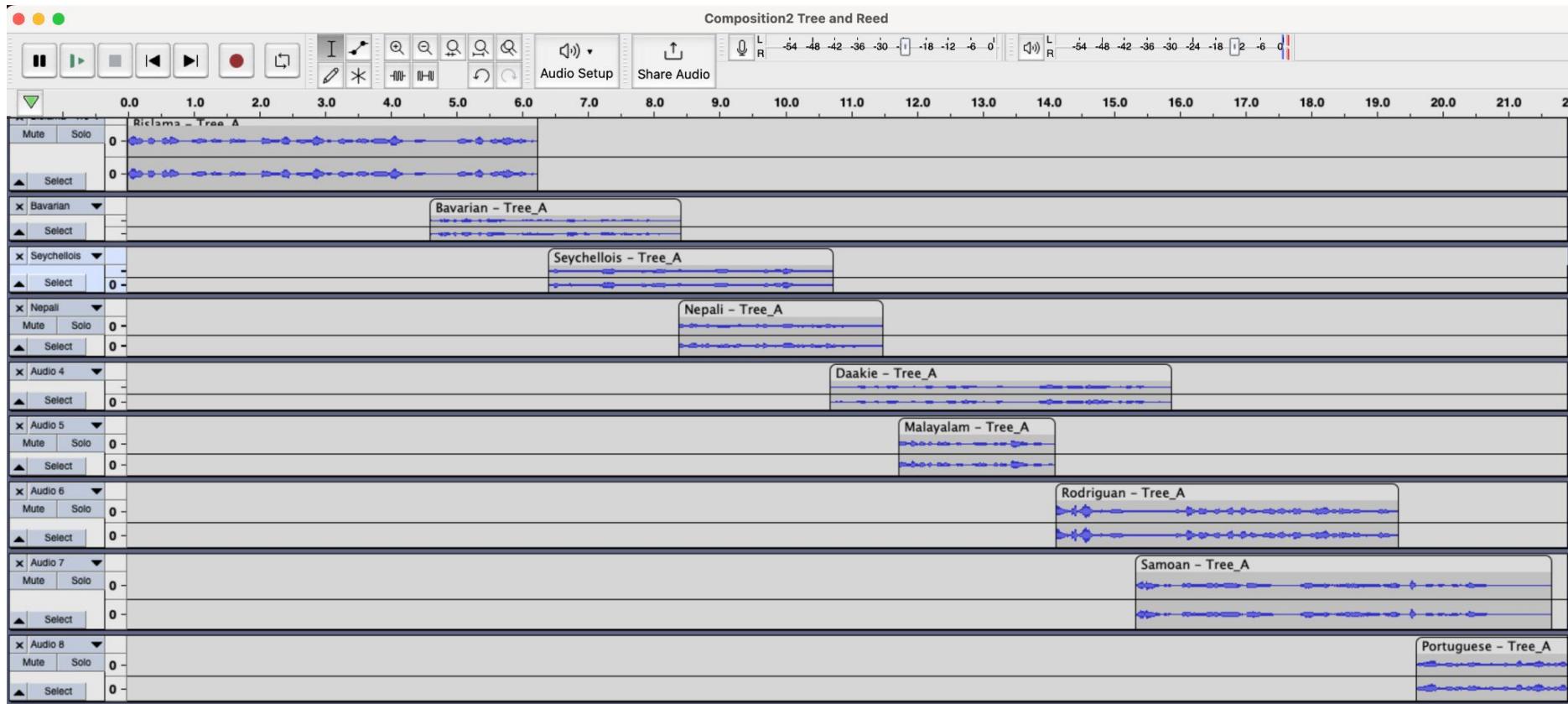
And finally the storm uprooted the mighty tree, and it fell down with a great **thunderous noise**.

The reed also felt the mighty storm, but it bent with the wind, this way and that way.

And when the storm was gone, the reed stood upright again.

The reed looked at the fallen tree, and said: It is better to yield when it would be unwise to resist than to resist stubbornly and be destroyed.

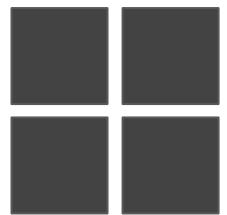
Realization of first two sentences



Future exciting goals

NMF - Strip FFT Bins to create a unique
Boquet of sounds for each speaker

icst plugins for spatial audio



icst plugins for spatial audio

LET IT BEE – TOWARDS NMF-INSPIRED AUDIO MOSAICING

Jonathan Driedger, Thomas Prätzlich, Meinard Müller

International Audio Laboratories Erlangen

{jonathan.driedger,thomas.praetzlich,meinard.mueller}@audiolabs-erlangen.de

ABSTRACT

A swarm of bees buzzing “Let it be” by the Beatles or the wind gently howling the romantic “Gute Nacht” by Schubert – these are examples of *audio mosaics* as we want to create them. Given a *target* and a *source* recording, the goal of audio mosaicing is to generate a *mosaic* recording that conveys musical aspects (like melody and rhythm) of the target, using sound components taken from the source. In this work, we propose a novel approach for automatically generating audio mosaics with the objective to preserve the source’s timbre in the mosaic. Inspired by algorithms for *non-negative matrix factorization* (NMF), our idea is to use update rules to learn an *activation matrix* that, when multiplied with the spectrogram of the source recording, resembles the spectrogram of the target recording. However, when applying the original NMF procedure, the resulting mosaic does not adequately reflect the source’s timbre. As our main technical contribution, we propose an extended set of update rules for the iterative learning procedure that supports the development of sparse diagonal structures in the activation matrix. We show how these structures better retain the source’s timbral characteristics in the resulting mosaic.

1. INTRODUCTION

Using the sounds in a recording of buzzing bees to recreate a recording of the song “Let it be” by the Beatles is a typical example of an audio mosaic. In this example, the recording of the bees serves as *source*, while the Beatles recording is called the *target*. Ultimately, one should be able to identify the target recording when listening to the mosaic, but at the same time perceive the timbre of the source sounds. Therefore, the audio mosaic of “Let it be” with the bee recording could give the impression of bees being musicians, buzzing the song’s tune.

Audio mosaicing is an interesting audio effect which has found its way into both artistic work as well as academic research. Artists like John Oswald used thousands of manually selected source audio snippets to create new

 © Jonathan Driedger, Thomas Prätzlich, Meinard Müller
Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Attribution: Jonathan Driedger, Thomas Prätzlich, Meinard Müller. "Let it Bee – Towards NMF-inspired Audio Mosaicing", 16th International Society for Music Information Retrieval Conference, 16th International Society for Music Information Retrieval Conference, 2015.

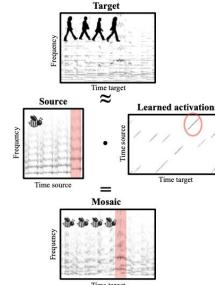


Figure 1. Schematic overview of our proposed audio mosaicing method. The sparse diagonal structures in the activation matrix are important in order to preserve the timbre of the source in the mosaic.

musical compositions¹ and real-time audio mosaicing has been used by musicians as an instrument in live performances [4, 22]. Over the years, many different systems for audio mosaicing were proposed [1, 3, 5, 11, 13, 17, 18]. The core idea of most automated systems is to split the source into short time segments, which are suitably concatenated afterwards to match spectral and temporal characteristics of the target [19].

In this work, we propose a novel way to create audio mosaics. Our idea is to learn an *activation matrix* that, when multiplied with the spectrogram of the source recording, approximates the spectrogram of the target recording (see Figure 1). The source spectrogram hereby serves as a *template matrix* which is fixed throughout the learning process. This way, as opposed to many previous automated mosaicing approaches, a frame of the target can be resynthesized as the superposition of several spectral frames of the source , thus allowing “polyphony” of the source sounds.

¹ Especially on his album *Pleasure* [16].

References

1. Gorenflo, LJ, S Romaine, RA Mittermeier, & K Walker-Painemilla. 2012. Co-occurrence of linguistic and biological diversity in biodiversity hotspots and high biodiversity wilderness areas. Proc Natl Acad Sci U S A 109(21), 8032-8037. <https://pubmed.ncbi.nlm.nih.gov/22566626>
2. Vilbjørg Broch. Spatial Audio Work - HOA. Frekvens Verden.
 - o Link: https://frekvensverden.dk/residency_iem_21.html
 - o Additional link: <https://frekvensverden.dk/spatial.html>
3. Gottschall, Jonathan. 2012. The storytelling animal. How stories make us human. New York: Houghton-Mifflin.
4. Franz Zotter & Matthias Frank. Ambisonics. Springer, 2019.
 - o Link: <https://link.springer.com/book/10.1007/978-3-030-17207-7>
5. Ambisonic Decoder Toolbox 2 (ADT2) - Python tool by Aaron Heller and Fernando Lopez-Lezcano, CCRMA, Stanford.
 - o Link: <https://bitbucket.org/ambidecodertoolbox/adt2/src/master/>

By Friday to add to this Presentation:

1 Slide: Etude (X?)

Different levels, whole story (mixing of sentences), sentence level (A-K), word level (400 files per language for example), syllable level (See Globalia), phoneme level

Mixing of these also possible

Vital levels for human communication

ToDo:

- For now output should be creating word level with Whisper AI () take this folder of files and give me 3 families (clusters)

- Apply various programs and see what the output is, think about the labeling.
- Words: Abkasian_A_4.wav
- Syllables
- Phonemes

Output: 1000s of files

Manually select certain pieces

Sonority hierarchy for stop hierarchy, use a universal classification of speech sounds.