



Lire à la volée le luxembourgeois : une innovation et un progrès pour l'accessibilité

ALAIN VAGNER
DOMINIQUE NAUROY



Retrouvez cette présentation en ligne

Les slides et leur transcription sont disponibles ici :

<https://gd.lu/5ZDRkh>





Qui sommes-nous ?



L'accessibilité numérique au Luxembourg : le Service information et presse



Accessibility Luxembourg

- Since 2020
- Surveillance de la conformité des sites et apps du secteur public
- Sensibilisation et formation
- Médiation entre les utilisateurs et les administrations





Qu'est-ce que l'accessibilité numérique ?



Un but

Garantir à tous, sans discrimination, le même accès à l'information et aux services en ligne, quelle que soit la nature du handicap des personnes et la façon dont chacun consulte l'information.



Pour faire face à des handicaps multiples



Comment naviguent les aveugles ?

Utilisation de **lecteurs d'écran** :
Vocalisation des contenus affichés à l'écran



Comment naviguent les aveugles ?

- Cette restitution
 - est immédiate
 - décrit le contenu et la structure d'une page,
 - décrit les fonctions de chaque élément interactif
 - fonctionne à très haute vitesse
 - s'adapte à toutes les langues présentes dans un même document
 - fonctionne offline





Acte 1

Le problème



Des outils qui manquent

- Pas de TTS luxembourgeois utilisable par les aveugles
- Pour contourner le problème, utilisation d'une synthèse vocale allemande
 - Frustration
 - Perte de temps
 - Charge cognitive supplémentaire



La situation... avant 2026

- Lecture d'un même article de rtl.lu par 3 synthèses vocales :
["Dat neit Loyersgesetz gött vu villen Immobilienagencen net respektéiert"](#)

- Extrait 1



- Extrait 2



- Extrait 3



Who's who?

- Extrait 1 – TTS luxembourgeois eSpeak
- Extrait 2 – TTS allemand macOS
- Extrait 3 – TTS luxembourgeois ZLS



Un projet qui s'apparente plutôt à un défi

- Interopérabilité
- Low-resource language
- Multilinguisme
- Open source





Acte 2 Wëllkomm Mia a Mil!



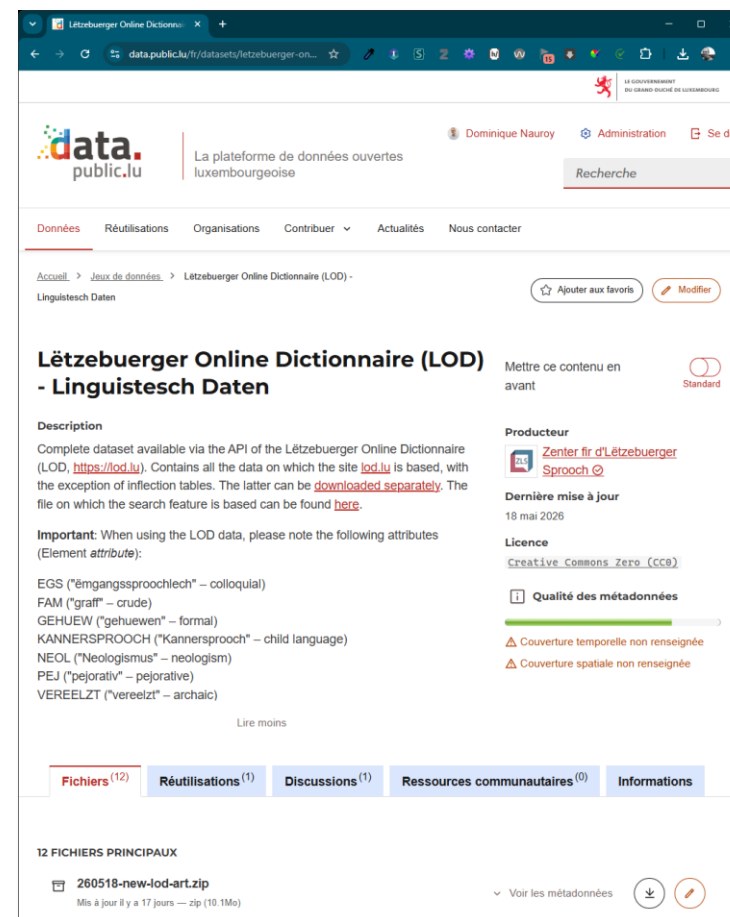
Les résultats de la solution développée

- Deux voix qui répondent au cahier des charges
 - Disponibles via RHVoice sur
 - iOS,
 - macOS,
 - Android,
 - Windows,
 - Linux



Qui s'est penché sur le berceau ?

- ZLS
 - Créateurs des données de base
 - Experts de la langue luxembourgeoise
 - Champions de l'open data
- CDV
 - Utilisateurs, testeurs
 - Experts accessibilité et usages
- Ministère de la Digitalisation
 - Financier



RHVoice, une solution sans réel challenger



RHVoice

- Compatible avec une majorité de lecteurs d'écrans et d'OS
- Compatible avec des « low power devices »
- Open source
- Expérience des « low resource languages »
 - 22 synthèses vocales dont le setswana, le tatar, le géorgien, le kirghize
 - Le luxembourgeois reste malgré tout un cas particulier



Qui a testé la solution ?



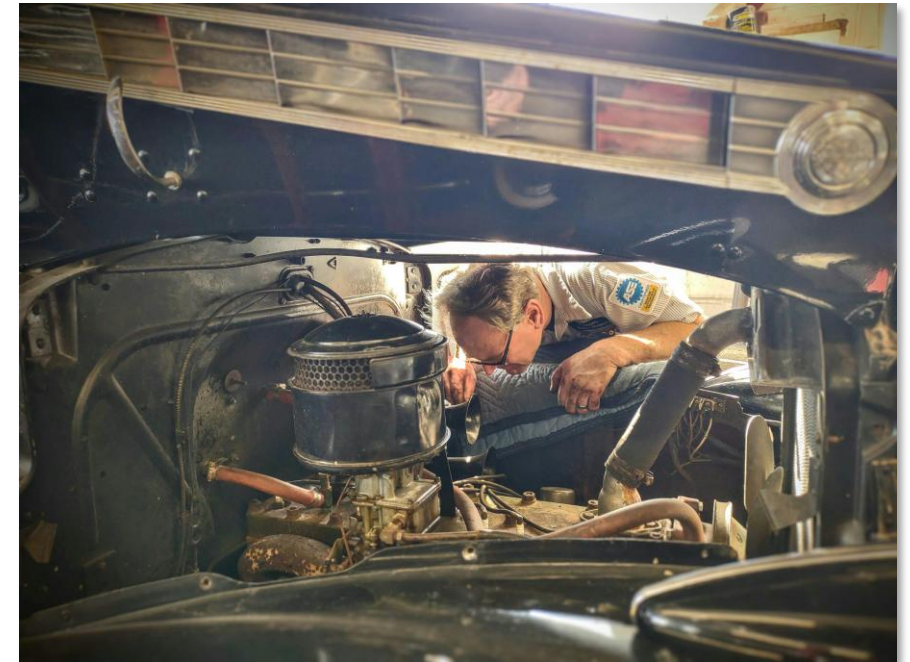
- Des personnes voyantes
 - Test en situation réelle
 - Tests synthétiques
- Des personnes aveugles
 - Tests en situation réelle, beta test publique
 - Organisation d'install parties





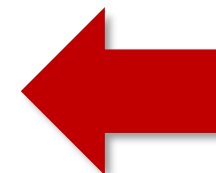
Acte 3

On ouvre le capot



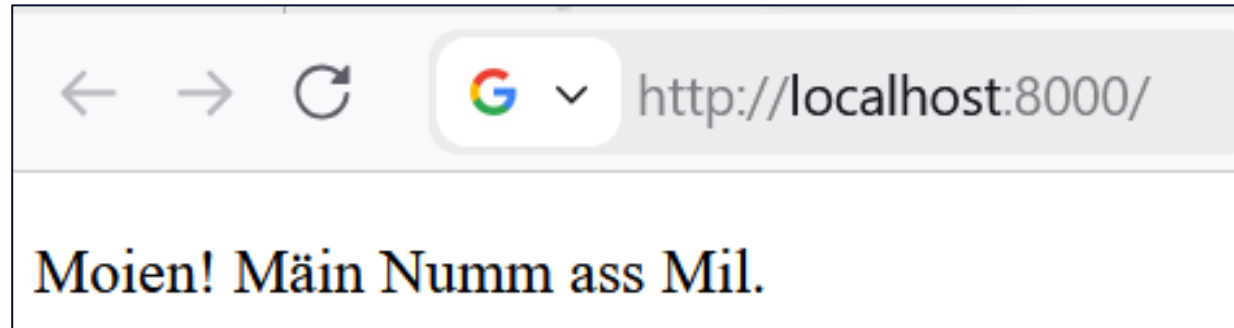
Les familles de synthèses vocales

Famille	Exemples de moteurs	Frugalité
Concaténative	Mbrola, Festival	★ ★
Paramétrique	eSpeak, RHVoice	★ ★ ★
Neuronale	Coqui TTS, Piper-TTS	★



TTS, le processus : 1. la source

Navigateur

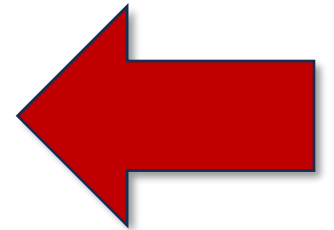


DOM HTML

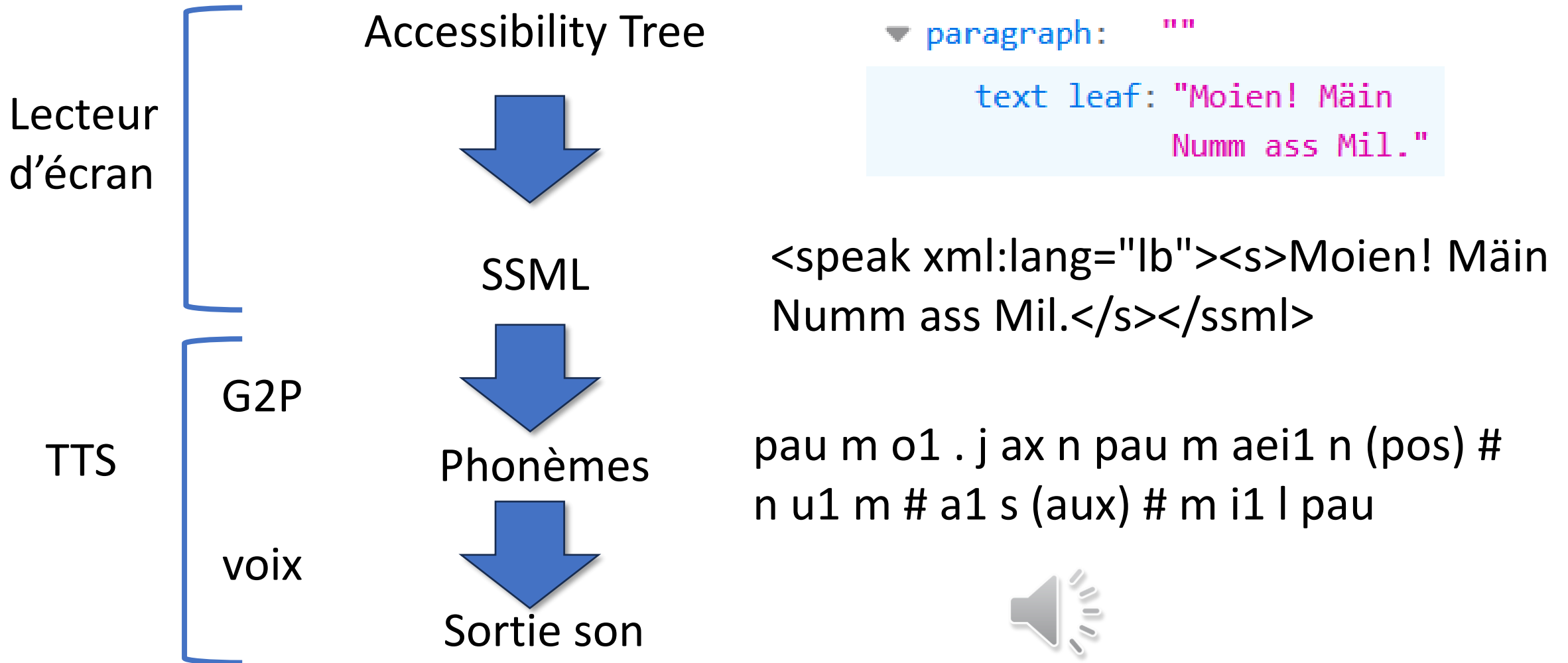
```
<p lang="lb">Moien! Mäin Numm ass Mil.</p>
```

Accessibility Tree

```
▼ paragraph: ""  
  text leaf: "Moien! Mäin  
              Numm ass Mil."
```



TTS, le processus : 2. la transformation



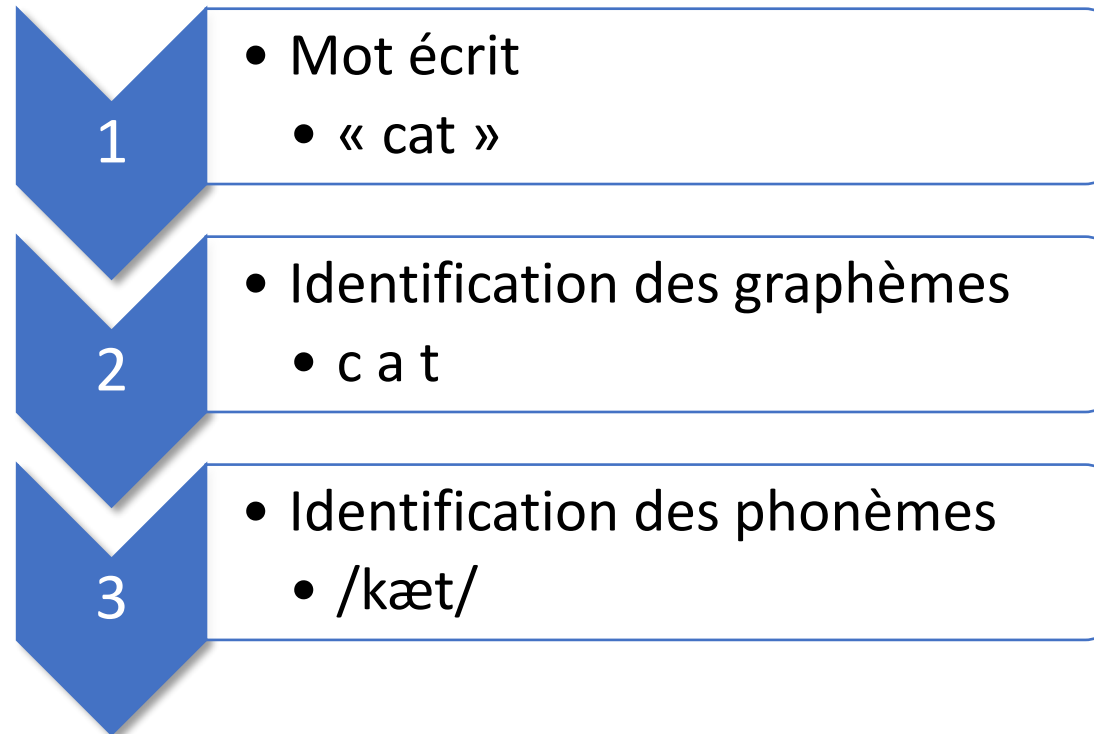
La voix, le modèle statistique paramétrique

- Sur base d'enregistrements fournis par le ZLS, apprentissage du mapping phonème -> son
- Boite noire
- Technologie HTS, Modèle de Markov caché



Deep Dive : le G2P, Grapheme to Phoneme

- G2P, kézako ?



Deep Dive : le G2P, Grapheme to Phoneme

- Basé sur le langage [FOMA](#)
 - Principe des « transducteurs » : [automates finis](#) qui produisent une sortie à partir d'une entrée
 - Transducteurs appliqués à la conversion graphème -> phonème
 - Code open source de notre G2P en FOMA :
<https://github.com/accessibility-luxembourg/RHVoice-Luxembourgish-src>



Deep Dive : le G2P, Grapheme to Phoneme

- Avec FOMA, description des règles de prononciation
- Règles
 - découpage de la phrase en mots,
 - prononciation de l'alphabet, des nombres,
 - gestion de mots dans le dictionnaire (prononciation phonétique),
 - gestion de mots non trouvés dans le dictionnaire,
 - mots composés,
 - symboles et emojis, abréviations,
 - intonations



Deep Dive : le G2P, Grapheme to Phoneme

Concrètement, à quoi ça ressemble, du code FOMA ?

Chat -> ʃa

Chien -> ʃjɛ̃

Crocodile -> kʁɔkɔdil

```
define Rules
c h -> ʃ
.o. c -> k || _ [a | o | u | r]
.o. i -> j || ʃ _
.o. e n -> ɛ̃
.o. o -> ɔ
.o. r -> ʁ
.o. e -> 0 || _ .# .
.o. t -> 0 || _ .# . ;
regex Rules;
```



Deep Dive : le G2P, Grapheme to Phoneme

En cascade : le processus G2P pour les mots luxembourgeois



1. Dictionnaire utilisateur
2. Dictionnaire luxembourgeois
3. Dictionnaire de français
4. Gestion des mots composés
5. Dernier recours : règles de prononciation génériques



Deep Dive : le G2P, Grapheme to Phoneme

- Cinq étapes pour le luxembourgeois... deux pour l'espagnol
- Les limites de cette approche
- Un même mot en luxembourgeois et français, deux prononciations différentes, ex: an, si, vu, ville, sou, rose



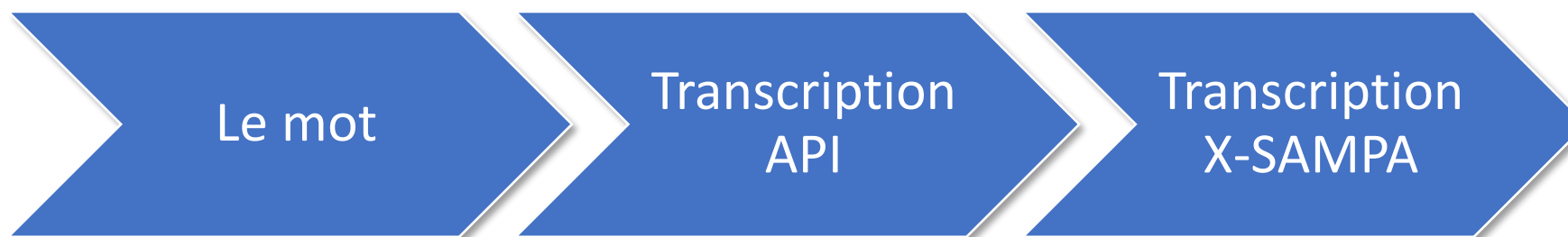
Deep Dive : le G2P, Grapheme to Phoneme

- Ajouter des mots au dictionnaire
 - Écoute des textes et détection des mots mal prononcés
 - Saisie phonétique avec l'outil interne
 - Vérification que la transcription X-SAMPA « compile »
 - Cycles d'écriture phonétique, test et écoute
 - Discussions pour les transcriptions de mots étrangers



Deep Dive : les dictionnaires

- Les dictionnaires et la phonétique



- **X-SAMPA** : notation développée pour transcrire l'**API** (Alphabet Phonétique International) en utilisant uniquement les caractères standards d'un clavier classique (ASCII 7 bits).
- Permet de représenter la prononciation des mots par une synthèse vocale.



Deep Dive : les dictionnaires

- Les dictionnaires et la phonétique : exemples

Mot	Transcription API	Transcription X-SAMPA	Détail des phonème
Maison	/mɛzõ/	"/mEzO~"	Voyelles et consonnes classiques + voyelle nasale (~)
Chat	/ʃa/	"/"Sa"	S pour le son "ch", a pour "a"
École	/ekɔl/	"/"ekOI"	O pour le son "o" ouvert
Thinking	/'θɪŋkɪŋ/	"/TINKIN"	T pour "th", I pour le "i" relâché, N pour le son "ng"



Deep Dive : les dictionnaires

- Le développement des dicos
 - 68 000 mots luxembourgeois
 - 244 000 mots français
 - Vocabulaire spécifique
 - Acronymes (ex: CGDIS, CTIE, ...)
 - Prénoms, noms de famille
 - Noms de sociétés
 - Célébrités, ...

Phonetic transcription tool for the Luxembourgish voices in RHVoice

Word Speak LOD X-SAMPA Voice

Consonants

b d D d̥ z dz f g h j k l m n N p R R\ s

s s\ t T t̥ s v w x z Z z\

Vowels

A a: a~ ɪ @ ɛ e E: e: E~ i i: o o: u u: y

Diphthongs

{:I { :U AI AU ɜI eI i@ ɛ_k oI oU @U u@

Tools

- [IPA to X-SAMPA converter](#)
- [IPA/Luxembourgish - Wikipedia \(source for most of the examples\)](#)



Deep Dive : les dictionnaires

- Le développement des dictionnaires : exemples

Mot	Transcription X-SAMPA
eenzeger	["e:nt_s@Z6]
nogelies	[n"og@li@s]
ponts	[p"o~n]
Jeannot	[Z"ano]
CePAS	["se:pAs]





Acte 4

Challenges & leçons



Les mots anglais

- Trop complexe d'intégrer une troisième langue
- Stratégie suivie : switch auto entre plusieurs voix



Les symboles et emojis

- Sauf exception, les symboles sont traités en amont de la synthèse vocale
 - Discussions avec les éditeurs de lecteurs d'écrans
- Une liste officielle des emojis est gérée par le [CLDR](#)
 - Traductions dans toutes les langues



Interopérabilité avec les lecteurs d'écrans

- Nombreux bugs sous iOS
 - Pas de contrôle sur les symboles et emojis
 - Symbole « . » lu « dot »
 - Certains boutons lu en anglais car LB pas reconnue
 - Certains textes envoyés au TTS sans majuscules



Interopérabilité avec les lecteurs d'écrans

```
<p lang="lb">Moien! Mäin Numm ass Mil.</p>  
<p lang="fr">Bonjour! Mon nom est Mil.</p>
```

Le support de l'attribut « lang » et le switch auto

- LB non reconnue sur iOS, pas de switch auto
- Sous Windows, switch auto dans les langues d'un moteur de synthèse
 - Problème: RHVoice ne supporte pas le français ou l'allemand



La logique de l'usage

- Innovation pensée dans un cadre « passif »
 - Lecture d'articles sur des pages web
 - Le principal usage semble être « actif »
 - Discussions WhatsApp, e-mails
- => Impact de l'orthographe sur la qualité de la restitution



La difficulté d'obtenir des feedbacks

- Processus d'adoption technologique
- Communication, gestion du changement, support
- Question de la rétribution des tests utilisateurs

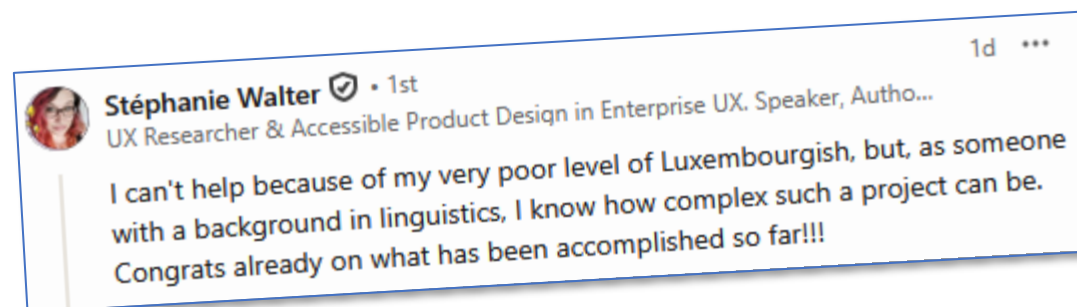
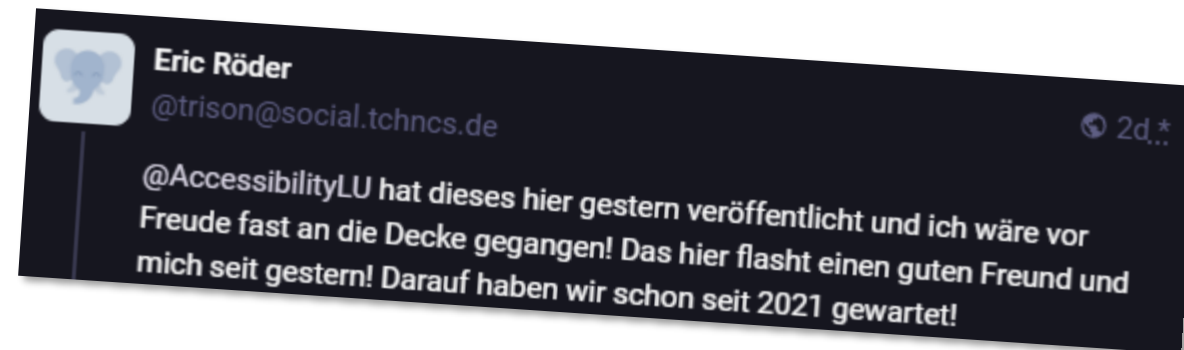
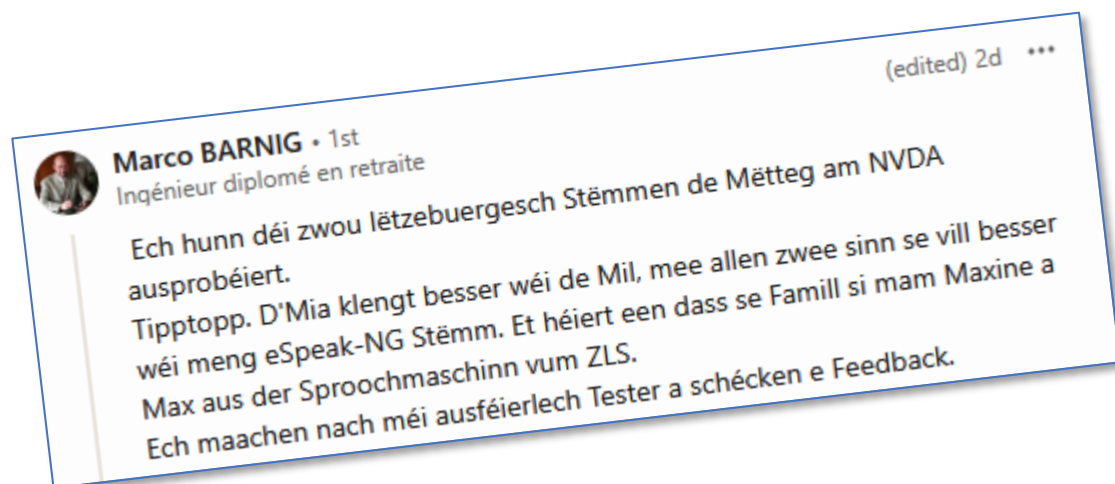


De premiers utilisateurs satisfaits

- **Trouvez-vous ces voix utiles ?** Oui, je les trouve utiles. Elles me facilitent la compréhension du luxembourgeois, contrairement à la version allemande.
- **Les utiliseriez-vous régulièrement ?** Oui, je pense.
- **Les recommanderiez-vous à vos collègues et amis ?** Oui, je les recommanderais car elles apportent une réelle valeur ajoutée.



Des retours positifs



Démo en ligne

Vous voulez tester les voix sans rien installer ?
Démo 100% dans le browser

RHVoice in the browser — Luxembourgish voices Mil & Mia

A proof of concept: the RHVoice synthesizer compiled to WebAssembly, reading Luxembourgish text and playing it back through the Web Audio API — no server.

Voice: Mil (male) ▾

Rate:  1.0x

 [Read this page](#)

<https://accessibility-luxembourg.github.io/rhvoice-emsripten-lb/>



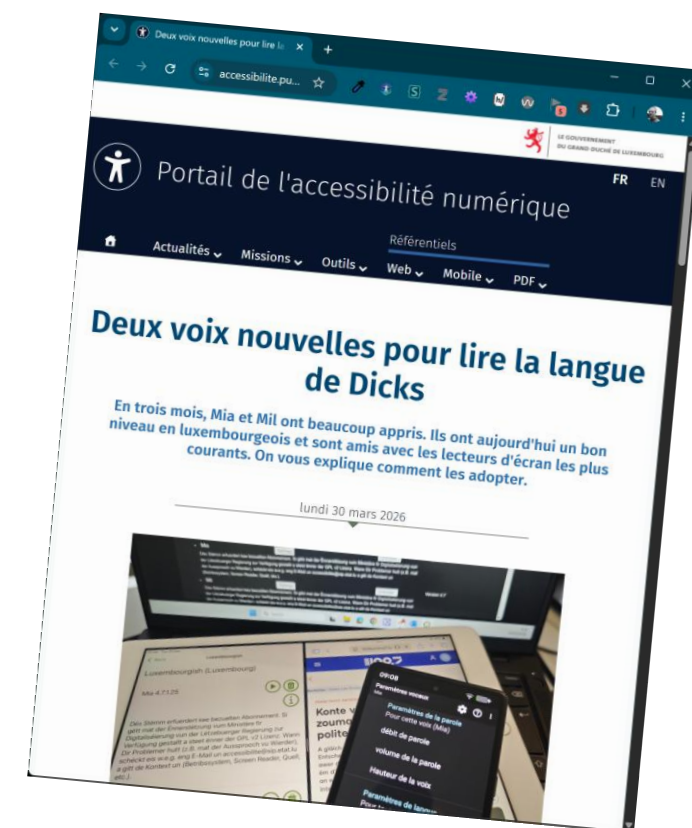


Conclusion



Un projet ambitieux, des résultats concrets

- Deux voix
 - offline,
 - cross platform,
 - open source,
 - réactives,
 - à l'aise même à haute vitesse
- Vous êtes invités à vous les approprier !
 - <https://accessibilite.public.lu/fr/news/2026-03-30-screenreaderLB.html>
- Contributions bienvenues !





Merci pour votre attention

Vos questions ? Nos réponses !

Nous contacter

accessibilite@sip.etat.lu

<https://accessibilite.public.lu>

Suivre nos actualités :

