

UiT
NORGES
ARKTISKE
UNIVERSITET

Publier en accès libre: comment structurer les données phonologiques?

Helene N. Andreassen
UiT Université arctique de Norvège

*Journées FLOrAL-PFC 2015: PFC dans le champ
phonologique*

3-8 décembre 2015, Maison de Norvège, Cité
Internationale Universitaire, Paris



Archivage en accès libre des données scientifiques

- Avantages de l'archivage
 - conserver
 - partager
 - citer
- Avantages de l'accès libre
 - découvrir
 - réutiliser
 - (ré)analyser



Exemple

Enchaînement, liaison, accentuation chez les apprenants norvégophones

Helene N. ANDREASSEN

UiT Université arctique de Norvège, BU Culture et Sciences Sociales & CASTL
Huginbakken 14, 9019 Tromsø, Norvège
helene.n.andreassen@uit.no

Chantal LYCHE

Université d'Oslo, ILOS
P. B. 1003, Blindern, 0315 Oslo, Norvège
chantal.lyche@ukis.uio.no

Given that their L1, Norwegian, and their L2, English, are lexical stress languages, Norwegian speakers will equally tend to stress lexical words during the course of acquisition of L3 French, insuring the prosodic autonomy of each word. In the present paper, we show that this strategy slows down the acquisition of two external sandhi phenomena in French, i.e. liaison and final consonant linking. The data, taken from two corpora recently collected in Tromsø and Oslo, indicate that the learners' acquisition path is conditioned by internal factors like prosodic weight, perceptual salience and frequency, and by external factors like the different tasks to be completed, i.e. reading vs conversation. The data further indicate that liaisons following determiners and clitics are the first categories to be acquired and that spontaneous speech, where the learner does not have direct access to the graphic word, seems to favor erasing the prosodic boundaries required by the prosodic system of her L1.

Keywords:

liaison, Consonant linking, prosody, external sandhi, Norwegian, French, L3 acquisition.

1. Introduction¹

La liaison, phénomène de sandhi externe particulier au français, consiste en l'apparition fortement conditionnée d'une consonne à l'initiale du mot-2. Si l'acquisition des modalités qui la régissent pose problème aux apprenants du français langue étrangère, l'autonomie prosodique du mot dans la L1 des apprenants constitue un obstacle supplémentaire, jusqu'à présent peu discuté dans la littérature. Ce travail a pour objectif principal d'examiner le chemin d'acquisition de la liaison àu sein de deux groupes d'apprenants norvégophones. Après une introduction à la prosodie du norvégien, nous présenterons et discuterons nos résultats à la lumière du phénomène



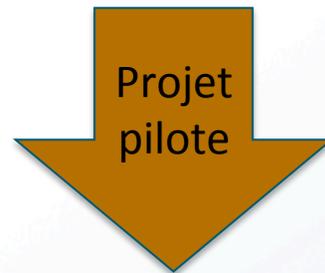
Investigation point	Informant code	Item	Category	Context	Target liaison consonant	Liaison co
Oslo	noosaf	nous	pro	avons	z	z
Oslo	noosaf	nous	pro	avons	z	z
Oslo	nooshi	nous	pro	avons	z	z
Oslo	noosjb	nous	pro	avons	z	z
Oslo	noosmh	nous	pro	avons	z	z
Oslo	noosms	nous	pro	avons	z	z
Oslo	noosaf	chemises	sub	chemises en soie	z	absence
Oslo	noosch	chemises	sub	chemises en soie	z	absence
Oslo	nooshi	chemises	sub	chemises en soie	z	absence
Oslo	noosjb	chemises	sub	chemises en soie	z	absence
Oslo	noosmh	chemises	sub	chemises en soie	z	absence
Oslo	noosms	chemises	sub	chemises en soie	z	absence
Oslo	noosaf	circuits	sub	circuits habituels	z	absence
Oslo	noosch	circuits	sub	circuits habituels	z	absence
Oslo	nooshi	circuits	sub	circuits habituels	z	absence
Oslo	noosjb	circuits	sub	circuits habituels	z	absence
Oslo	noosmh	circuits	sub	circuits habituels	z	absence
Oslo	noosms	circuits	sub	circuits habituels	z	absence
Oslo	noosaf	dans	pre	dans une	z	z
Oslo	noosch	dans	pre	dans une	z	z
Oslo	nooshi	dans	pre	dans une	z	z
Oslo	noosjb	dans	pre	dans une	z	z
Oslo	noosmh	dans	pre	dans une	z	z
Oslo	noosms	dans	pre	dans une	z	z
Oslo	noosaf	des	det	des activistes	z	z
Oslo	noosch	des	det	des activistes	z	z

Archivage de données linguistiques



UiT Open Research Data

Share, Cite, Reuse, Archive Research Data
Scientific data for reproducible research



[Opendata.uit.no](https://opendata.uit.no)

Instructions disponibles en ligne

The TROLLing Blog



The screenshot shows the homepage of the TROLLing Blog. At the top, there is a navigation bar with links for Home, About, TROLLing archive, Getting started, Promotional materials, Presentations, People, and Contact. Below the navigation bar is a header section with a blue background and a white text box containing the title "TROLLing" and the subtitle "The Tromsø Repository of Language and Linguistics". To the left of the text is a small image of a boat on water. Below the header is a search bar with a "Search" button. The main content area is divided into several sections: "Getting started" with a sub-section "User guides:" containing two links to PDF guides, "Instruction videos:" containing four links to YouTube videos, "Recent Posts" with a list of five recent posts, and "Recent Comments" and "Archives" sections.

Getting started

User guides:

- [TROLLing essentials](#) (pdf)
- [TROLLing data format essentials](#) (pdf)

Instruction videos:

- [How to create an account](#) (YouTube)
- [How to create a study](#) (YouTube)
- [How to upload data](#) (YouTube)
- [How to search for data](#) (YouTube)

Recent Posts

- [TROLLing data format essentials](#)
- [TROLLing at Dataverse Community Meeting](#)
- [TROLLing at emtacl 15](#)
- [Poljarnyj vestnik requests archiving in TROLLing](#)
- [TROLLing at the Journées FLORAL in Paris](#)

Recent Comments

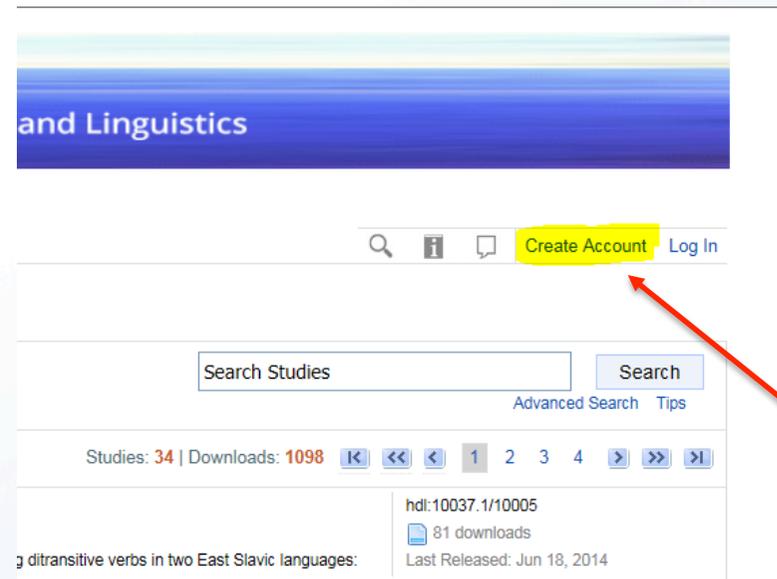
Archives

- [November 2015](#)
- [June 2015](#)
- [April 2015](#)

<http://site.uit.no/trolling/getting-started/>

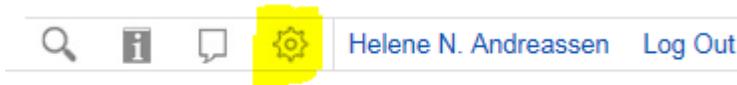
Etape 1: Créer un profil dans TROLLing

- Tout le monde peut accéder les données dans TROLLing, sans nom d'utilisateur/mot de passe
- Uniquement des personnes ayant le statut de *TROLLing contributor* peuvent verser des données dans l'archive
- Vous créez un profil, les administrateurs de TROLLing vérifient que vous êtes une vraie personne (!) et vous donnent ensuite les droits nécessaires

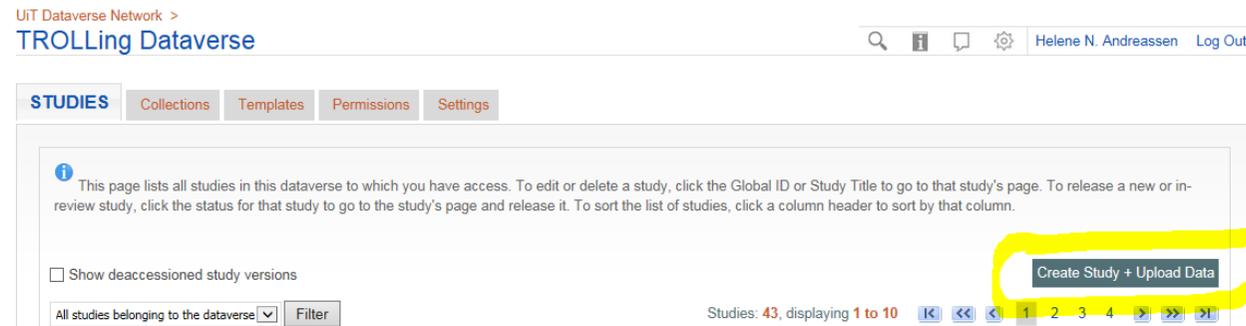


Etape 2: Création d'une étude (« study » = dataset)

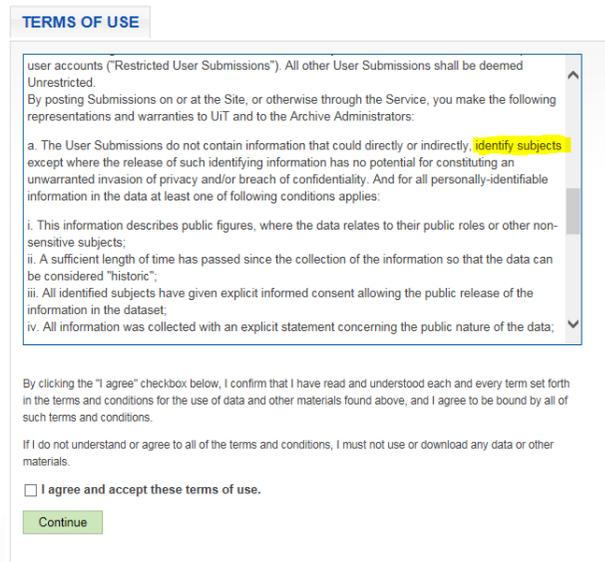
1



2



3



Vous êtes seul responsable des données. Vous garantissez que vous avez l'autorisation de publier les données, qui doivent impérativement être anonymisées



Getting started

User guides:

- [TROLLing essentials \(pdf\)](#)
- [TROLLing data format essentials \(pdf\)](#)

Etape 3: Enregistrer des métadonnées

Data Citation Details

The Data Citation is generated automatically by the Dataverse based on the information entered below.

Title *

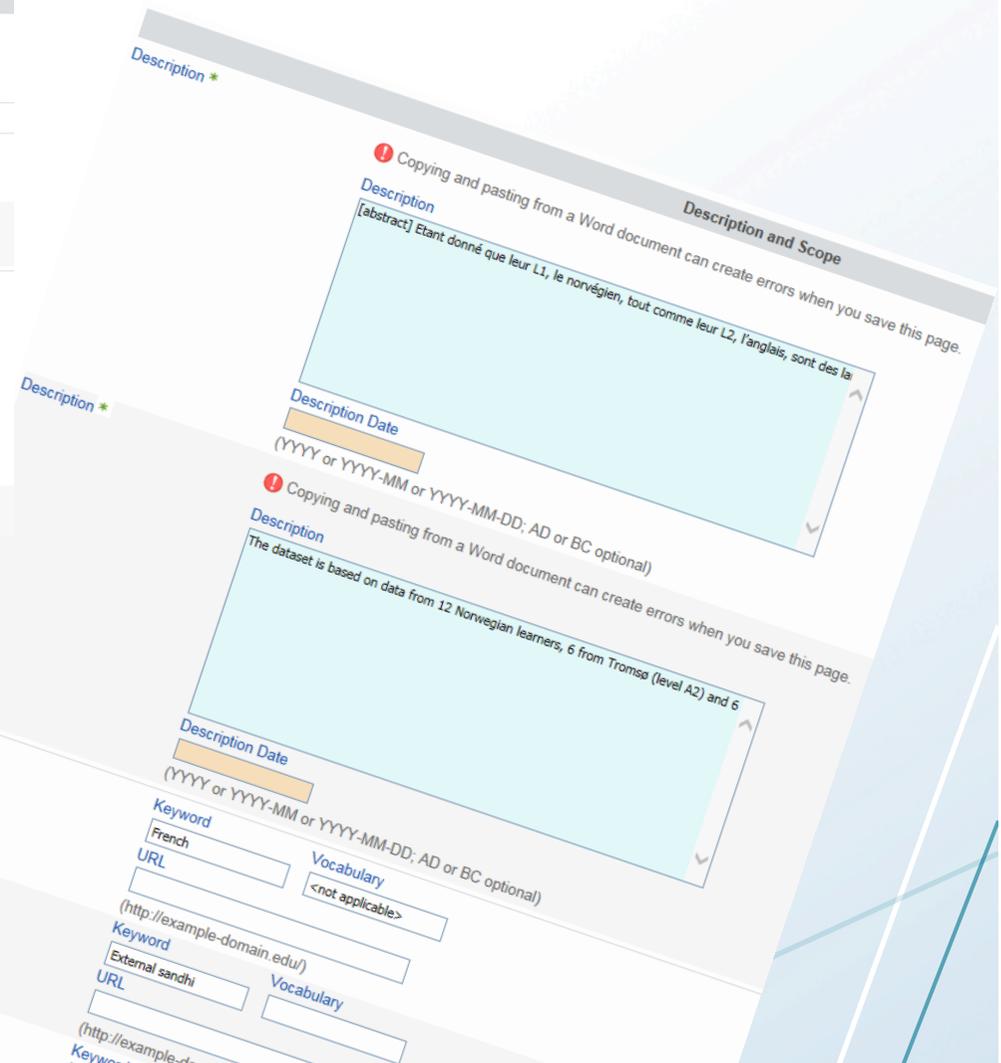
Study Global ID

Author *

Name	<input type="text" value="Andressen, Helene N."/>	Affiliation	<input type="text" value="UIT The Arctic University of Norway"/>
<small>(LastName, FirstName)</small>			
Name	<input type="text" value="Lyche, Chantal"/>	Affiliation	<input type="text" value="University of Oslo"/>
<small>(LastName, FirstName)</small>			

Producer *

Producer	<input type="text" value="UIT The Arctic University of Tromsø"/>	Abbreviation	<input type="text" value="UIT"/>
Affiliation	<input type="text"/>		
URL	<input type="text" value="http://www.uit.no"/>		
<small>(http://example-domain.edu/)</small>			
Logo URL	<input type="text"/>		
<small>(http://example-domain.edu/image)</small>			
Producer	<input type="text" value="University of Oslo"/>	Abbreviation	<input type="text" value="UIO"/>
Affiliation	<input type="text"/>		





Getting started

User guides:

- [TROLLing essentials \(pdf\)](#)
- [TROLLing data format essentials \(pdf\)](#)

Etape 4: Verser des données

Format des fichiers

- Tableaux Excel
 - format .txt
- Fichiers son
 - wav, flac, mp3
- Transcriptions
 - Unicode ou SAMPA
 - TextGrid en combinaison avec csv
 - Word en combinaison avec pdf

File type	Acceptable file formats	Unacceptable file formats (examples)
Audio	<ul style="list-style-type: none"> • Uncompressed and lossless Wav or AIFF (.wav/.aiff) • Compressed and lossless FLAC (.flac) • Compressed and lossy Mp3 (.mp3) 	<ul style="list-style-type: none"> • AAC (.m4a) • Monkey's Audio (.ape) • Ogg Vorbis (.ogg) • Windows Media Audio (.wma)
Image	<ul style="list-style-type: none"> • Uncompressed TIFF (.tif or .tiff) • Compressed and lossless PNG (.png) • Compressed and lossy JPEG (.jpg) 	<ul style="list-style-type: none"> • Adobe Photoshop (.psd) • Apple Picture File (.pct) • Graphics Interchange Format (.gif) • Raw Image Data File (.raw) • Windows Bitmap (.bmp)
Text (slides)	<ul style="list-style-type: none"> • PDF/A (.pdf) combined with original file 	<ul style="list-style-type: none"> • PowerPoint (.pptx)
Text (tables)	<ul style="list-style-type: none"> • PDF/A (.pdf) combined with Comma/Tab Separated Values (.csv/.txt) 	<ul style="list-style-type: none"> • Excel (.xlsx)
Text (text)	<ul style="list-style-type: none"> • Plain text (.txt) • If formatting needed • XML, PDF/A (.pdf) combined with original file 	<ul style="list-style-type: none"> • Word (.docx) • HTML
Transcription	<p>File format</p> <ul style="list-style-type: none"> • PDF/A (.pdf) combined with original file • PDF/A (.pdf) combined with Comma/Tab Separated Values (.csv/.txt) <p>Font</p> <ul style="list-style-type: none"> • Unicode IPA (e.g. Charis SIL, Doulos SIL, Gentium Plus, Andika), ASCII SAMPA 	<p>File format</p> <ul style="list-style-type: none"> • Word (.docx) • Excel (.xlsx) <p>Font</p> <ul style="list-style-type: none"> • Transcription legacy fonts (SIL IPA(93))
Video	<ul style="list-style-type: none"> • MPEG-4 (.mp4) 	<ul style="list-style-type: none"> • AVI (.avi) • Flash Video (FLV) • Quicktime (.mov) • Windows Media Video (WMV)

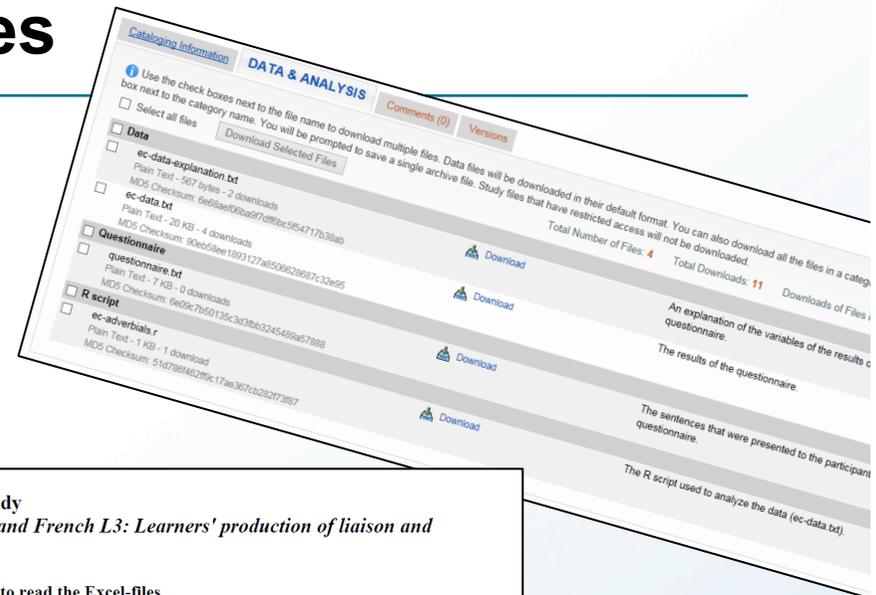
Etape 4: Verser des données

Catégoriser

- Regrouper les fichiers par type de fichier ou locuteur

Expliquer

- Expliquer le contenu de chaque fichier, soit dans la colonne à droite, soit dans un fichier séparé (*Read me* file)
- Tableau Excel: pour aider d'autres chercheurs, ou vous-même, à lire (et réutiliser) vos données, vous devez expliquer chaque colonne



TROLLing study Norwegian L1 and French L3: Learners' production of liaison and enchaînement

User guide: How to read the Excel-files

- 1) The consonants are phonetically transcribed using the SAMPA alphabet, and there are thereby two differences compared to the standard IPA alphabet
 - a. /Z/ is used for /ʒ/
 - b. /S/ is used for /ʃ/
- 2) Data sorting order: Investigation point – (register –) Context – Informant code
- 3) Column information
 - a. *Investigation point* = informants' origin/place where the recordings took place: Oslo, Tromsø. Note that in the respective corpora, all informants were born and grew up in the same place.
 - b. *Informant code* = anonymous and unique code given to each informant participating in the IPFC project: no = Norway; os/tr = Oslo/Tromsø; xx = initials of first and last name; number = number to disambiguate

Investigation point	Informant code	Register	Item	Category	Context	Target liaison consonant	Liaison consonant
Oslo	noosmh	free	trop	adv	agréable	p	absence
Oslo	noosjb	free	trop	adv	amusant	p	absence
Oslo	noosjb	free	appris	ver	appris un	z	z
Oslo	noosch	free	as	aux	as été	z	absence
Oslo	nooshi	free	as	aux	as étudié	z	absence
Oslo	noosjb	free	as	ver	as une	z	absence
Oslo	noosjb	free	avait	aux	avait eu	t	absence
Oslo	noosch	free	beaucoup	adv	beaucoup à	p	absence
Oslo	noosmh	free	beaucoup	adv	beaucoup appris	p	absence
Oslo	noosaf	free	c'est	cop	c'est aussi	t	absence
Oslo	nooshi	free	c'est	cop	c'est aussi	t	absence
Oslo	noosjb	free	c'est	cop	c'est en	t	t
Oslo	noosch	free	c'est	cop	c'est important	t	absence
Oslo	noosch	free	c'est	cop	c'est important	t	absence
Oslo	noosmh	free	c'est	cop	c'est intéressant	t	absence
Oslo	noosmh	free	c'est	cop	c'est intéressant	t	t

Etape 5: Publication des données

SUBMIT FOR REVIEW

Ensuite

- Possibilité de modifier les métadonnées et les données
- Création automatique d'une nouvelle version
- La référence bibliographique indique la version actuelle

Version History					
Difference	Version	Status	Comments	Released	Contributors
<input type="checkbox"/>	5	Released		Wed Jun 18 10:31:09 CEST 2014	Helene N. Andreassen
<input type="checkbox"/>	4	Archived		Mon May 26 13:41:12 CEST 2014	Helene N. Andreassen
<input type="checkbox"/>	3	Archived		Wed May 21 11:52:48 CEST 2014	Helene N. Andreassen
<input type="checkbox"/>	2	Archived		Tue May 20 10:41:45 CEST 2014	Helene N. Andreassen
<input type="checkbox"/>	1	Archived		Tue May 20 10:18:54 CEST 2014	Helene N. Andreassen

Data Citation

i If you use these data, please add the following citation to your scholarly references. [Why cite?](#)

Mykhaylyk, Roksolana; Rodina, Yulia; Anderssen, Merete, 2014, "Replication data for: Ditransitive constructions in Russian and Ukrainian", <http://hdl.handle.net/10037.1/10005> UiT Open Research Data [Distributor] V5 [Version]

Citation Format [Print](#) ▾

Original Publication

i Results found in this publication can be replicated using these data.

Mykhaylyk, R., Y. Rodina & M. Anderssen. 2013. Ditransitive constructions in Russian and Ukrainian: Effect of givenness on word order. *Lingua*, 137, 271-289.

[Link](#)

UiT
NORGES
ARKTISKE
UNIVERSITET

**Pour toute question, n'hésitez
pas à nous contacter**

Helene

helene.n.andreassen@uit.no

The TROLLing team

trolling@ub.uit.no

