PAMAPLA 22

Papers from the 22nd Annual Meeting

of the

Atlantic Provinces Linguistic Association

University College of Cape Breton
Sydney, Nova Scotia, Canada
6-7 November 1998

ACALPA 22

Actes du 22ième Colloque annuel

de l'Association de linguistque des Provinces atlantiques

Collège universitaire du Cap-Breton
Sydney, Nouvelle-Écosse, Canada
les 6-7 novembre 1998

Edited by / Rédaction

William Davey and Bernard LeVert

PAMAPLA 22

Papers from the 22nd Annual Meeting

of the

Atlantic Provinces Linguistic Association

University College of Cape Breton
Sydney, Nova Scotia, Canada
6-7 November 1998

ACALPA 22

Actes du 22ième Colloque annuel

de l'Association de linguistque des Provinces atlantiques

Collège universitaire du Cap-Breton
Sydney, Nouvelle-Écosse, Canada
les 6-7 novembre 1998

Edited by / Rédaction

William Davey and Bernard LeVert

Printed in 12 point New Times Roman and bound by / Impressions et reliure: Print Shop, University College of Cape Breton, Sydney, Nova Scotia, Canada

Papers from the 22nd Annual Meeting of the Atlantic Provinces Linguistic Association
Actes du 22ième Colloque annuel de l'Association de linguistique des Provinces atlantiques

Legal Deposit / Dépôt légal: 1998
National Library of Canada
Bibliothèque nationale du Canada

ISBN: 0831-3520

APLA / ALPA

22nd Annual Meeting / 22ième colloque annuel

Organizing Committee / Comité d'organisation

Stephanie Inglis
Bernard LeVert
William Davey

Acknowledgments / Remerciements

The Atlantic Provinces Linguistic Association thanks the following for their contribution to the 22nd Annual Conference and the publication of this volume.

L'Association de linguistique des Provinces atlantiques tient à remercier les personnes et organismes suivants pour leur appui financier et moral.

Dr. Thomas Rendall, Dean of the School of Arts and Letters

Dr. Robert Morgan, Executive Director of Research and Development

Department of Languages and Letters

TABLE OF CONTENTS / TABLE DES MATIÈRES

List of Other Papers Presented / Liste d'autres communications présentées

Raimo Anttila (UCLA)
Ironing out Modern Variation with the Stone Age: English ACKER/ACRE

Katherine Barber (Oxford University Press)
Establishing a Standard for Canadian English; and
Making Dictionaries is not Dull Work: Workshop on Dictionary Making in a Canadian Context

Louise Beaulieu (Université de Moncton, Campus de Shippagan) and
Wladyslaw Cichocki (University of New Brunswick)
The Role of Social Network in Analyzing *que* Variation in an Acadian French Speech Community

Catherine Bodin (Western Maryland College)
Spécialisation lexicale dans le français cadien louisianais: cinq exemples

W. Terrence Gordon (Dalhousie University)
Boiling, Hanging, Driving, Learning: Observations on the Morphology, Syntax, and Semantics of C.K. Ogden's BASIC English

John Hewson (Memorial University of Newfoundland)
SVO and the Autonomy of Syntax

Ruth King (York University) and **Terry Nadasdi** (University of Alberta)
On Variable Use of Traditional Acadian Verbal Morphology in Prince Edward Island French

Martyna Kozlowska-Macgregor (McGill University)
Semantic Operations within the Syntax of Polish Aspect

Donna L. Lillian (York University)
Modality, Persuasion and Propaganda

Raymond Mopoho (Dalhousie University)
Le français africain: entre imitation, rejet et appropriation

Lewis J. Poteet (Concordia University)
From Offiside to Offline: Recent Evolution in Business Slang

Authors' Addresses / Adresses des auteur-e-s

Raimo Anittila
1914 1/2 Euclid St.
University of California
Santa Monica, CA90404-4610
anttila@ucla.edu

Katherine P.Barber
Oxford University Press
70 Wynford Drive
Toronto, ON M3C 1J9
kpb@oupcan.mail.net

Louise Beaulieu
Bureau du vice-recteur
Université de Moncton
Centre universitaire de Shippagan
Shippagan, N-B EOB 2PO
louise@cus.ca

Catherine Bodin
Western Maryland College
2 College Hill
Westminster, MD 21157-4390
USA
cbodin@msmary.edu

Wendy Burnett
Department of Modern Languages
and Literatures
49A York Street
Mount Allison University
Sackville, NB E4L 1C7
wburnett@mta.ca

Walcir Cardoso
1804 Tupper, apt. #1
Montréal, PQ
Canada H3H 1N4
wcardo@po-box.mcgill.ca

Wladyslaw Cichocki
Department of French
University of New Brunswick
Fredericton, NB E3B 5A3
cicho@unb.ca

Sheila Embleton
Department of Languages, Literatures
and Linguistics
South 929 Ross Building
York University
4700 Keele Street
Toronto, ON M3J 1P3
embleton@yorku.ca

Michael Falk
1591 Conrose Avenue
Halifax, NS B3H 4C4
falk@fox.nstn.ca

W. Terrence Gordon
Department of French
Dalhousie University
Halifax, NS B3H 3J5

John Hewson
Department of Linguistics
Memorial University of Newfoundland
St. John's, NF A1B 3X9
jhewson@morgan.ucs.mun.ca

Ruth King
Department of Languages, Literatures
and Linguistics
York University
4700 Keele Street
Toronto, ON M3J 1P3
rking@yorku.ca

Martyna Kozlowska-Macgregor
McGill University
Department of Linguistics
1001 Sherbrooke Street West
Montréal, PQ H3A 1G5

Donna L. Lillian
Department of Languages, Literatures
and Linguistics
York University
4700 Keele Street
Toronto, ON M3J 1P3
dlillian@yorku.ca

Cheng Luo
Department of Applied Language Studies
Brock University
St. Catharines, ON L2S 3A1
cluo@spartan.ac.brocku.ca

Raymond Mopoho
Department of French
Dalhousie University
Halifax, NS B3H 3J5
rmopoho@is.dal.ca

Terry Nadasdi
Department of Modern Languages
University of Alberta
Edmonton, AL T6G 2E2
tnadasdi@gpu.srv.ualberta.ca

Agnès Picolet-Crépault
Université Sainte-Anne
Pointe-de-l'Église
Nouvelle-Écosse BOW 1MO
picolet@ustanne.ednet.ns.ca

Lewis J. Poteet
604 les Érables
Laval sur le lac, Québec H7R 1B6
Lpoteet@vax2.concordia.ca

Une comparaison du parler acadien traditionnel avec le parler populaire dans l'œuvre de Molière

Wendy Burnett
Mount Allison University

RÉSUMÉ

Grâce à des études importantes sur le français acadien, qui ont le but de décrire soit les origines de ce parler, soit son système linguistique, soit ses variantes lexicales, nous savons que plusieurs formes du français archaïque et populaire du 17ième siècle se sont maintenues jusqu'au 20ième siècle dans le parler acadien traditionnel. Cette analyse se propose de montrer les correspondances phonétiques, morphologiques, syntaxiques, et lexicales qui existent entre le parler populaire de l'Ile-de-France tel qu'il est présenté dans l'œuvre de Molière, et le parler acadien traditionnel. Deux questions secondaires se posent au cours de l'analyse. On se demande, d'abord, si le dramaturge connaissait bien le parler populaire de son époque. La deuxième question porte sur l'exactitude de la reproduction du parler populaire par l'écrivain et sa sensibilité aux nuances sociolinguistiques.

Introduction

Plusieurs études ont été publiées sur le français acadien, dans le but de décrire soit les origines de ce parler (Poirier, Massignon), soit son système linguistique (Lucci, Ryan, et Flikeid entre autres), soit les deux à la fois (Geddes, Péronnet, Gérin & Gérin). Grâce à ces auteurs, qui ont fait du parler acadien un domaine particulier, nous savons aujourd'hui que plusieurs formes du français archaïque et populaire du 17^{e} siècle, et surtout du Centre-Ouest de la France, se sont maintenues jusqu'au 20^{e} siècle dans le parler acadien traditionnel.

Captivée par les nombreux commentaires sur les ressemblances entre le parler acadien et le parler des paysans dans le *Dom Juan* de Molière, j'ai entrepris un projet de recherche qui a eu l'ambition de faire une étude systématique des correspondances phonologiques, morphologiques, syntaxiques , et lexicales qui existent entre le parler populaire de l'Ile-de-France tel qu'il est présenté dans les comédies de Molière, et le parler acadien traditionnel.

Il est important de souligner qu'on se sert du terme 'le parler acadien' là où il s'agit des traits communs aux parlers acadiens de toutes régions.

Parmi les 31 comédies de Molière, j'en ai choisi 22 pour l'étude, une sélection basée sur le nombre de termes populaires, archaïques, ou vieillis y figurant. Logiquement, les termes de l'enquête, et surtout l'étude phonétique, exigeaient que la recherche soit executée sur des reproductions des éditions originales des pièces.

Un certain nombre de questions secondaires se sont posées au cours de l'analyse. Il fallait se demander, d'abord, si Molière connaissait bien le parler populaire de son époque. La deuxième question portait sur l'exactitude de la reproduction du parler populaire et la sensibilité de l'écrivain aux nuances sociolinguistiques. Une dernière question s'est aussi posée: quel effet Molière recherchait-il en employant le langage familier? S'il ne cherchait qu'à faire rire son public, le parler de ses paysans ne serait-il qu'une collection de traits conventionnels, pittoresques et comiques, modelé sur le parler paysan employé dans les pièces comiques du 16ième siècle? Nous pensons, par exemple, aux pièces de Cyrano de Bergerac.

Les critiques de *Dom Juan* et du *Médecin Malgré-Luy*, les deux pièces dans lesquelles le parler populaire est souvent commenté, s'accordent sur la première question: le parler dont se servent les paysans est 'le langage du menu peuple de la ville de Paris et des paysans de la banlieue parisienne' (Lagerqvist 1974:366).

Le contexte sociolinguistique

Quant aux nuances sociolinguistiques, une étude minutieuse de toutes les comédies où apparaît le parler populaire a révélé que Molière était remarquablement au fait de la variation sociale et situationnelle.

En premier lieu, j'ai pu constater que Molière n'a pas créé de rôles paysans pour les farces qu'il a écrites pendant son séjour en province. Est-ce parce que la bourgeoisie, public par excellence des tournées en province, était trop habituée au parler populaire des domestiques et des paysans pour le trouver comique? Certes, on stigmatise peu souvent le parler que l'on entend tous les jours.

En deuxième lieu, il paraît que, sans les connaître, Molière suit les principes de William Labov. Ses personnages féminins emploient moins de formes stigmatisées que les hommes et sont plus sensibles qu'eux aux modèles de prestige. Les conventions de la variation situationnelle (où le locuteur dispose du style qui correspond à la situation dans laquelle il se trouve) demandent que dans *Dom Juan*, Charlotte, qui patoise toujours moins que son fiancé, le paysan Pierrot, utilise un niveau de langue encore plus élevé en s'adressant à Dom Juan. Le dialectologue Dauzat souligne que des observateurs impartiaux au 17^{e} siècle ont noté que les jeunes paysannes considéraient leur patois comme une langue inférieure dont elles avaient plus ou moins honte (Dauzat 1927:30).

L'étude des traits phonétiques

L'étude phonétique est composée d'une analyse détaillée de toutes les variantes relevées dans les 22 comédies étudiées et indiquées dans les tableaux comparatifs. Le premier tableau présente les correspondances vocaliques. Les exemples en lettres se trouvent dans les comédies; ceux en symboles phonétiques proviennent

du parler acadien traditionnel.

Tableau comparatif 1: les variantes vocaliques
FS=français standard; V=variante; PA= parler acadien

FS	V	Exemple	PA	Molière
i	ɛ	*vaignes*		▫
e	i	*agriable*		▫
e	a	*acoute*		▫
ɛ	e	*teste*	▫	▫
ɛ	œ	*aveuc, queul*	▫	▫
ɛr	ar	*Piarrot*	▫	▫
ɔ	u	*estoumaque*	▫	▫
ɔ	œ	*quement*	▫	▫
o	u	*roole*	▫	▫
o final	jo	*biaux*	▫	▫
u	o	[kotym]	▫	
u	ɔ	[ʒɔrne]	▫	
ʏ	i	*ribans*	▫	▫
ø final	u	*vieilloux*	▫	▫
œr final	ør	[sør]	▫	
les voyelles nasales				
ɛ̃	ɛɲ	*chagraignes*	▫	▫
ɛ̃	on	[fon]'faim'(N-E)	▫	
ɛ̃	ã	*rian*		▫
ɔ̃	ã	*an* 'on'	▫	▫
œ̃	ɛ̃	[saʃɛ̃]'chacun'	▫	
les semi-consonnes				
wa	wɛ	*boeste*	▫	▫
wa	ɛj	*nayé*	▫	▫
wa	u	[sune]'soigner'	▫	
war final	we	[mirwe]'miroir'	▫	

Il faut d'abord remarquer que malgré ses talents prodigieux pour l'interprétation des niveaux de langue de ses personnages, Molière ne disposait que de l'alphabet latin pour exprimer les dialectes. Sans avoir recours aux symboles phonétiques, il ne pouvait distinguer entre le /ɑ/ du FS et le phonème dialectal /ɑ/ qui s'articule au même niveau que /ɔ/ et parfois encore plus en arrière (Lucci 1972: 63). Il ne pouvait différencier non plus l'ouverture d'une voyelle devant les occlusives finales, ni la fermeture des voyelles accentuées.

Malgré certaines différences entre les parlers populaires étudiés, surtout dans les voyelles nasales, il y a des correspondances frappantes. Il me semble important de souligner que plusieurs traits distinctifs du parler acadien existent dans le parler populaire de Molière, dont les plus significatifs sont le remplacement du /ɛr/ par /ar/, comme dans la réalisation «vart» (MML), la voyelle nasale /ɔ̃ / réalisée par /ã/

dans «*sant*» (DJ, MML, FS), et le phénomène du 'ouisme' qui donne [hum] 'homme' et [pum] 'pomme'. Il n'y a que deux exemples du 'ouisme' dans les 22 comédies, notamment «*estoumaque*» (DJ II, 1) et «*roole*» (PE I, 4); étant donné les maintes possibilités pour la manifestation de cette variante dans le corpus, il se peut que l'écart n'ait pas été dominant dans la région de Paris au 17ième siècle.

D'importance secondaire sont la confusion de /ɔ/ et /œ/ en position inaccentuée, comme dans «*quement*» (DJ II, 1); le /u/ pour le /ø/ final dans des adjectifs comme «*vieilloux*» (DJ II, 1); les deux variantes de la semi-consonne /wa/ dans les exemples «*coeffeure*» (EM 1, 1) et «*sayent nayez*» (DJ II, 1) et la palatalisation de la voyelle /o/ dans plusieurs pièces: *gliau* l'eau', *biau* 'beau', *la Biausse* 'la Beauce'.

On trouve 8 variantes dans les parlers acadiens qui ne se trouvent pas dans les parlers populaires des textes littéraires: /u/ prononcé [o] *'godron'*, et [ɔ] *'jornée'*; /ɛ/ prononcé [œ], comme dans *'achoeve'*; /ɛ̃/ prononcé [on], par exemple [pon] 'pain'; /œ̃/ se dit [ɛ̃] dans les mots comme [kœk zɛ̃] 'quelques-uns' ; et /wa/ prononcé [u] comme dans [sune] 'soigner', un trait propre à la Nouvelle-Ecosse. Quatre écarts caractéristiques du parler populaire de l'Ile-de-France au 17ième siècle (i>ɛ: *'chopaine'*; e>i: *'agriable'*; e>a: *'acoute'*; ɛ>a: *'bian, rian'*) se voient souvent dans le parler moliéresque, mais ne se trouvent pas dans les parlers acadiens traditionnels.

C'est au plan consonantique qu'on trouve les plus grandes divergences entre les parlers acadiens et les parlers populaires de Molière.

Tableau comparatif 2: Les variantes consonantiques

FS	V	Exemple	PA	Molière
la palatalisation				
k	g	*segret*	□	□
d	j	[djø]	□	
tj/dj	kj/gj	*quieu, amiquié*	□	□
g	dj/dʒ	[djɛp]'guêpe'	□	
ʃ	ʒ	[ʒval]	□	
ʒ	dj	*adjouster*	□	□
ʒ	χ	[χamɛ]	□	
ʒ	h	[hamɛ]	□	
les formes affriquées				
t	tʃ	[tʃɛ̃d]	□	
ʃ	tʃ	[tʃãsɔ̃]	□	
ʒ	dʒ	[dʒur]	□	
le mouvement postérieur				
p	b	*Biarre, béres*		□
l	r	*pal* 'par'	□	□
l	n	[menas]'mélasse'	□	

On voit par ce tableau qu'en comparaison avec les réalisations du parler acadien traditionnel, les phénomènes de palatalisation sont beaucoup moins complexes dans les parlers populaires de Molière, où, en général, le processus implique la palatalisation des consonnes dentales suivies du yod, par exemple dans «*piquié* pour 'pitié, et «*heriquié* pour 'héritié'. Il est intéressant que la palatalisation des consonnes vélaires en /dj/ ou /ʤ/ ne se trouve pas dans les textes étudiés; pourtant ce phénomène remonte au 15e siècle (Gérin & Gérin 1982:184) et s'est révélé assez durable. En revanche, il n'est peut-être pas surprenant que les consonnes affriquées n'existent pas dans les textes littéraires. Selon les études de Péronnet, c'est un trait particulier à la région de l'Ouest (Péronnet 1989:248).

Les procédés phonétiques

Le parler populaire dans l'œuvre de Molière partage peu des procédés phonétiques qu'on rencontre dans le parler acadien traditionnel. En ce qui concerne la syncope, on voit tomber le /r/ et le /l/ suivant une autre consonne, par exemple dans «*putost*» et «*notte*». Par contre, la chute du /v/ initial devant /wa/ dans l'acadien, par exemple dans [wɛjaʒ] 'voyage' n'existe pas dans les comédies. L'apocope, très général dans le parler acadien traditionnel: [avɛ] 'avec', [ʃeti] 'chétif', [avri] 'avril', est rare dans le parler populaire des comédies, à l'exception de la chute du /r/ final, comme dans les mots «*toujou*» et «*leus*». L'épenthèse ne paraît que dans «*estoumaque*» et l'adjectif «*vilainte*». La métathèse aussi ne se voit que rarement: *une éplingue* 'épingle' (DJ II, 1).

Les variantes morphologiques

On voit par le Tableau comparatif 3. que la plus grande ressemblance entre les parlers acadiens traditionnels et le parler populaire moliéresque se trouve au plan morphologique.

Le rapport entre les deux parlers est particulièrement proche dans le cas des pronoms personnels, où la plupart des variantes sont du type morphophonologique: 'elle' prononcé [al]; la réalisation [i] pour le pronom sujet pluriel 'ils'; la forme «*li*» pour 'lui'; la chute du 'r' dans 'leur'. D'ordre morphophonologique aussi sont les variantes du pronom indéfini 'quelque', prononcé [kœ/øk] dans les deux parlers. Les formes des pronoms démonstratifs se rapprochent assez étroitement de celles du parler acadien, surtout dans le cas de «*cety-ci, cety-la*» (MML II,10) 'celui-ci, celui-là'.

Les terminaisons à la 3e personne du pluriel *-ont, -ant* pour le temps présent, et *-iont,-iant* pour l'imparfait, le présent du subjonctif et le conditionnel sont des formes dialectales conservées dans le parler populaire, marquant l'opposition

(Gérin & Gérin 1982:134). Grevisse affirme leur emploi assez général dans les dialectes d'oïl du célèbre [ʒə] 'nous', un emploi qui apparaît au 15^{e} siècle.

Tableau comparatif 3: Les variantes morphologiques

FS	V	Exemple	PA	Molière
le pronom personnel				
elle	*al*		□	□
ils	*i*		□	□
lui	*li*		□	□
leur	*leu*		□	□
eux	[zø]		□	
nous	*je*	*j'avions*	□	□
le pronom indéfini				
quelque [ø/œ]		*queuque chose*	□	□
le pronom démonstratif				
formes semblables		*cety-ci, cety-la*	□	□
les formes verbales				
-ent	-ont/-ant	*is ant*	□	□
-aient	-iont/-iant	*ils faisiant*	□	□
verbe intrans. avec 'avoir'		[ʒe tõbe]	□	
verbe pronom. avec 'avoir'		*i s'a mis*		□
usage du passé surcomposé			□	□
le nom				
écarts de genre		*par la sang*	□	□
/al/o/ final		[ʃval]'chevaux'	□	
l'adjectif				
[t] comme marque du sing.		[ɛ̃ grot um]	□	
forme féminine avec [t]		*vilainte*	□	□

Les écarts grammaticaux des parlers acadiens de provenance régionale se trouvent rarement dans les parlers de Molière. La réalisation *'al'* pour le pronom sujet 'elle', qui se voit dans les deux groupes, provient des régions de Normandie et du Poitou, mais, selon Péronnet, la forme s'est étendue à Paris et ses environs (Péronnet 1995: 420).

Bien que l'usage de l'auxiliaire 'avoir' avec les verbes intransitifs perfectifs soit courant dans le parler acadien traditionnel, comme dans la plupart des parlers d'oïl en France, le patois de la région de Paris employait 'être' au 17^{e} siècle, si

Molière est exact dans sa représentation. Cette possibilité est attestée par la carte 1312 de l'ALF. Péronnet constate que l'usage de l'auxiliaire 'avoir' en France est très peu répandu avec les verbes pronominaux (Péronnet 1991:92); les formes notées dans le parler populaire des comédies de Molière correspondent seulement au parler acadien lousianais. Ditchy cite: «*il s'a laissé tomber, je m'ai trompé*». Pour ce qui est de l'usage du passé surcomposé dans le parler d'un campagnard: «*je n'avais pas putost eü gagé que . . .* »(DJ II, 1), nous trouvons un parallèle dans le parler acadien du sud-est du Nouveau-Brunswick.

L'adjectif interrogatif 'quel' est réalisé [kœ] [kjœ]ou [ʧœ] dans le parler acadien traditionnel, mais se dit «*queul*» dans Molière, aussi bien que dans l'acadien traditionnel louisianais. '*Zeux*', originaire du Centre-Ouest, n'apparaît pas dans Molière. On ne voit non plus la liaison 't' qui marque le singulier, comme dans '*un grot homme*'.Cependant, on rencontre l'adjectif dont la forme féminine a un 't', comme dans le cas de '*vilainte*': «*t'es une vilainte, toy*»(DJ III, 1), qui correspond aux formes relevées par Péronnet: [kryt]'crue', [purit]'pourrie' (Péronnet 1989:80).

Les variantes syntaxiques

Le parler populaire dans l'oeuvre de Molière s'écarte rarement du français standard sur le plan syntaxique. Là où la variation existe, Molière a tendance à adopter les écarts qui proviennent du français archaïque plutôt que de la langue populaire, comme par exemple: l'usage fréquent de la préposition 'à' comme marque d'appartenance, un 'que' ajouté à la préposition 'avant de' dans une expression telle que «*avant que de vous la faire voir*», la conservation de quelques conjonctions archaïques comme 'à cause que'.

Tableau comparatif 4. Les variantes syntaxiques

FS	V	Exemple	PA	Molière
le complément de relation				
de	à	*à un turc*	□	□
le pronom relatif				
dont, lequel	que		□	
que	où			□
l'inversion				
interrogative			rare	□
non-interrogative		*ce ma til fait*	□	□
redondance des pr. interrogatifs			□	
la particule *-ti*				
avec l'interrogatif			□	
usage non-interrogatif			□	□

Le parler populaire des comédies de Molière ne partage pas la redondance des pronoms interrogatifs qu'on entend dans le parler acadien: *qui ce qui?*, *qui c'est qui?* , etc. On n' y voit pas non plus les manifestations de l'invariabilité communes aux parlers populaires, par exemple, l'invariabilité du genre des noms et des adjectifs, l'ordre sujet-verbe pour l'interrogatif, l'usage de l'indicatif pour le subjonctif, et les expressions de volonté suivies du conditionnel. Cependant, les deux parlers partagent les expressions intensives des parlers populaires, comme le pléonasme: «*par fin pour faire court*» (DJ II, 1), «*mon père est toujours malade de plus en plus*» (MP I, 6), la suppression de 'ne' et l'usage de 'point' avec la négation, le redoublement du sujet: «*je ne sçay pas moy*» (MML I, 4), «*nous autres*» (GD I, 1), et les transformations analogiques, comme les termes médicaux déformés du *Médecin Malgré-Luy*, tels que «*infection*» pour 'infusion', «*conversion*» pour 'convulsion', et «*syncole*» pour 'syncope'.

L'écart le plus frappant, à mon avis, est l'usage de -ti comme forme non-interrogative, par exemple dans le cas de «*ce ma til fait*» pour 'il m'a dit', une forme qui ressemble fortement à l'usage de la particule -ti dans les parlers acadiens.

Les variantes lexicales

Les variantes lexicales relevées dans l'oeuvre de Molière sont au nombre de 70, un nombre trop élevé pour dresser un tableau comparatif. Parmi ces 70 mots archaïques ou populaires, 26 sont mentionnés dans les glossaires de Boudreau, Maillet, Massignon, Péronnet et Poirier. Si le pourcentage de correspondances, qui s'élève à 37%, ne semble pas significatif à première vue,
il ne faut pas oublier que les informateurs et les répondants des études sur les parlers acadiens sont issus d'un milieu rural, et que les locuteurs du parler populaire dans les comédies, au contraire, sont présentés pour la plupart dans un contexte urbain.

Conclusion

On peut conclure que l'analyse phonétique, grammaticale, et lexicale des deux parlers indique des correspondances étroites entre le parler populaire de l'Ile-de-France, tel que présenté dans les comédies de Molière, et le parler acadien traditionnel.

Au commencement de mon étude, j'avais voulu déterminer si Molière connaissait le Poitou et les parlers poitevins. Il était bien probable que Molière et sa troupe étaient passés par cette région au cours de leurs douze années de tournées en province, mais il n'était pas facile d'en établir les preuves. Après avoir fouillé longtemps, j'ai découvert un document attestant la présence de la troupe Dufresne-Molière à Poitiers, le premier novembre 1648, par la mention dans les archives de la paroisse Saint-Cybard des obsèques de la femme de Dufresne.

Marcel Cohen se demande d'où Molière tenait l'emploi de «*de même*» pour 'pareil, ainsi', puisque cet usage n'était pas encore attesté dans la région de Paris à son époque. Cohen suggère qu'il le devait à l'influence de l'Ouest, et comme preuve il relève l'usage dans le Supplément aux *Glossaires du Poitou* (Cohen 1951:85), témoignage linguistique qui sert d'appui au document mentionné ci-dessus. Encore mieux, la variante «*vilainte*» constitue une autre preuve. Selon Dauzat, cet écart n'existe pas dans les textes patois de la région parisienne, et a son origine en Saintonge et en Poitou occidental (Dauzat 1927:142).

Pourquoi Molière a introduit ces deux variantes nous ne saurons jamais, mais il semble qu'il connaissait le parler du Poitou septentrional, et avait remarqué ses particularités en comparaison avec le parler populaire de la région de Paris.

BIBLIOGRAPHIE

BOUDREAU, Ephrem. 1988. *Glossaire du vieux parler acadien: mots et expressions recueilli à Rivière-Bourgeois* . Montréal: Éditions du Fleuve.

COHEN, Marcel. 1951. "L'adjectif invariable 'de mëme' en Poitou et dans Molière" dans *Français moderne*, XIX, pp. 87-89.

DAUZAT, Albert. 1927. *Les Patois*. Paris: Librairie Delagrave.

DITCHY, Jay. 1932. *Les Acadiens louisianais et leur parler*. Réimpression de l'édition de Paris: Librairie Larousse.

GÉRIN, Pierre et Pierre Gérin. 1982. *Marichette: Lettres acadiennes 1985-1898.* Sherbrooke: Editions Naaman.

LAGERQVIST, Hans Äke. 1974. "Une forme inattendue dans le patois de Dom Juan." dans *Moderna Sprak*, LXVII, pp. 366-372.

LUCCI, Vincent. 1972. *Phonologie de l'acadien*. Montréal: Didier.

MAILLET, Antonine. 1980. *Rabelais et les traditions populaires en Acadie*. Québec: Les Presses de l'Université Laval.

MASSIGNON, Geneviève. 1962. *Les Parlers français d'Acadie*. Tomes I&II. Paris: Librairie C. Klincksieck.

Œuvres complètes de Molière. Tomes I-XI. 1947. Texte établi et annoté par Gustave Michaut. Paris: Imprimerie nationale.
Cités: *Dom Juan* (DJ), *L'Escole des maris* (EM), *Les Fourberies de Scapin* (FS), *George Dandin* (GD), *Le Médecin Malgré-Luy* (MML), *Monsieur de Pouceaugnac* (MP), *La Princesse d'Élide* (PE) .

PÉRONNET, Louise. 1989. *Le parler acadien du Sud-Est du Nouveau-Brunswick*. New York: Peter Lang.

---------- . 1991. "Système des modalités verbales dans le parler acadien du Sud-Est du Nouveau-Brunswick" dans *Revue de l'Association de Linguistique des Provinces Atlantiques*. Vol. 13, pp. 85-98.

---------- . 1995. "Le Français acadien." dans *Français de France et français du Canada*. Lyon: Centre d'études linguistiques Jacques Goudet, p. 399-439.

POIRIER, Pascal. 1928. *Le parler franco-acadien et ses origines*. Québec: Imprimerie Franciscaine missionaire.

---------- . 1993. *Le Glossaire acadien*. Édition critique établie par Pierre M.Gérin. Moncton: Centres d'études acadiennes.

Variation Patterns In Regressive Assimilation In Picard
An Optimality Theoretic Account*

Walcir Cardoso

McGill University & Universidade Federal do Pará

0 Introduction:

This paper offers an Optimality Theoretic account for the phonological process of Across-Word Regressive Assimilation (AWRA henceforth) in Vimeu Picard, a Gallo-Romance dialect spoken in the Picardie region in Northern France. In my investigation, I focus on one particular topic in the analysis of AWRA: the variation patterns observed in the application of AWRA; the analysis is couched within the framework of Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1993ab). The data come from Vasseur (1963) and Debrie's (1981) articles on Picard, as well as from Auger's database consisting of a fieldwork sample of 2,783 tokens of variants of AWRA; the latter were collected following a three-level stylistic interview with nine speakers of Picard in five villages of Vimeu: Nibas, Feuquières, Fressenneville, Bienfay and Bouillancourt.

The investigation presents an account of variation using the tools of Optimality Theory (OT henceforth). Following the works of Reynolds (1994), Nagy and Reynolds (1997), Anttila (1997) and Taler (1997), the quantitative analysis that I present for AWRA in Picard provides support for the view that variation can be encoded in the grammar through variable ranking of constraints. In this way, my analysis uses tools from theoretical linguistics in the form of Optimality Theory and includes quantitative findings to provide an account of the non-categorical regularities found in the AWRA process in language use. This study is thus an attempt "to account for the broadest possible range of facts about language, including usage as well as abstract knowledge" (Guy 1997) within a single (competence) grammar (cf. Labov 1969 et al., Wardhaugh 1994, Guy 1997, among several other variationist linguists). In other words, the study attempts to incorporate into the grammar much more than what is usually referred to as such, i.e. aspects of the use of language.

This paper is composed of three main sections. In section 1, I provide the data that illustrate AWRA and its variation patterns, and demonstrate that the phenomenon is exclusively sensitive to the phonological phrase domain. Section 2 sets forth my motivations for the selection of the frameworks of Optimality Theory and Prosodic Phonology to analyze the AWRA data. Finally, in section 3, I introduce the subject of variation in Optimality Theory, illustrate the results of the quantitative analysis and propose an OT analysis for the variation patterns involved in AWRA.

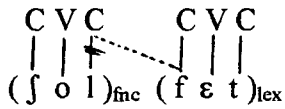
* I am extremely grateful to Heather Goad for her guidance and encouragement throughout the development of this research. For their insightful comments, I would also like to thank the audience of the Twenty-second Annual Meeting of the Atlantic Provinces Linguistic Association. This research has also benefited from e-mail conversations with Arto Anttila and Keren Rice, and helpful input from Julie Auger (who first introduced me to the Picard world). Finally, I would like to acknowledge funding from a FCAR grant to Julie Auger (NC-1648). I alone am responsible for any errors that remain.

1 The data: Prosodic domain and variation

The relevant data upon which my study is based are reviewed in this section. I will first provide evidence that the AWRA phenomenon is sensitive to prosodic domains inasmuch as the scope of the process is limited to a specific prosodic configuration. Second, I illustrate the variation patterns encountered in clitic plus lexical word sequences in Picard.

As I show in (01), AWRA is a domain-sensitive phonological phenomenon that operates exclusively at the domain juncture of a (C)Vl shape clitic (l-clitic henceforth, *fnc* in (01)), followed by a consonant-initial lexical word (*lex*). When both phonological and morphosyntactic contexts are met, the Root node of the lexical word's initial consonant associates to the timing slot of the previous clitic-final consonant /l/, resulting in a complete geminate across the two words:[1]

(01) The AWRA process:



→ [ʃof fɛt] 'the party'

In (02), (03) and (04), I demonstrate that the application of AWRA is sensitive to the prosodic domains in which the constituents involved prosodicize (i.e. a clitic followed by a lexical word), as well as to the melodic properties of the following segment (i.e. consonant vs. vowel). Observe in (2a) through (2f) that the clitic-final consonant /l/ completely assimilates to the following onset when the l-clitic (which is by definition unstressed) is followed by a consonant-initial lexical word, regardless of its grammatical status (i.e. Noun, Adjective, etc.).

(02) AWRA contexts (Vasseur 1963, Debrie 1981 and field notes of Auger and my own):

a.	/ ʃol glɛn /	→	[ʃog **gl**ɛn]	'the / this chicken' (Determiner ʃol)
	/ ʃol vak/	→	[ʃov **v**ak]	'the / this cow'
	/ ʃol pjøt tab/	→	[ʃop **pj**øt tab]	'the / this small table'
b.	/ dol grɛs /	→	[dog **gr**ɛs]	'some fat' (Partitive dol)
	/ dol vjan /	→	[dov **vj**ãn]	'some meat'
	/ dol bɔn vjan /	→	[dob **b**ɔn vjãn]	'some good meat'
c.	/ al fɛt /	→	[**af** fɛt]	'at the party' (Preposition al)
	/ al kaʃ /	→	[**ak** kaʃ]	'at the hunting'
	/ al mem plaʃ /	→	[**am** m̃ɛm ...]	'in the same place'
d.	/ al dãs /	→	[**ad** dãs]	'she dances' (3rd Person al)
	/ al va/	→	[**av** va]	'she goes / is going'
e.	/ ɛl fis /	→	[ɛ**f** fis]	'the son' (Determiner ɛl)
	/ ɛl mɔn /	→	[ɛ**m** m̃ɔ̃n]	'the world'

[1] The representation in (01), in particular the CV tier, is used for illustrative purposes only. As you will see in section 3, I will dispense with the CV tier in favor of prosodic constituents.

f. / i ɛl pɔrt /	→	[i **p** pɔrt][2]	'He brings it'	(Object pronoun ɛl)
/ va ɛl vir /	→	[va **v** vir]	'is going to see it'	

Nevertheless, AWRA does not apply in phonological and syntactic contexts distinct from the ones illustrated in (02). Observe below that /l/-faithfulness (or inapplicability of AWRA) is the result if the following lexical word is vowel-initial, as in (03), or when the relevant sequence of consonants occurs in monomorphemic words, in prefixation, in compounding and in other higher syntactic configurations; see (4a-d).

(03) Inapplicability of AWRA (phonological environment):

a. Vowel-initial word:	/ ʃol armwɛr / →	[ʃ **l** armwɛr][3]	'the closet'
	/ dol arb / →	[d **l** arb]	'the tree'

(04) Inapplicability of AWRA (morphosyntactic environment):

a. Monomorphemic word:	/ kalfa / →	[kal**f**a], *[ff]	'caulker'
	/ belʒik/ →	[belʒik], *[ʒʒ]	'Belgium'
b. Prefixation:	/ malpoli / →	[mal**p**oli], *[pp]	'impolite'
	/ malʃãs/ →	[malʃãs], *[ʃʃ]	'bad luck'
c. Compounding:	/ belmer / →	[bel**m**er], *[mm]	'mother-in-law'
	/ belfij / →	[bel**f**ij], *[ff]	'daughter-in-law'
d. Other domains:	/ bel tab / →	[bel **t**ab], *[tt]	'beautiful table'
	/ drol dø / →	[drol **d**ø], *[dd]	'funny (of)'

Based on these facts, I conclude that an analysis that recognizes the interaction between phonology and syntax (i.e. Prosodic Phonology) is necessary for a comprehensive account of the AWRA process in Picard. An analysis which does not refer to prosodic domains predicts the illicit forms in (04). Cardoso (1998a) argues that the patterns found in /l/-clitic plus lexical word sequences is governed by the prosodic status that each word assumes in the grammar. More specifically, he argues that the AWRA phenomenon applies exclusively at the domain juncture of an unstressed

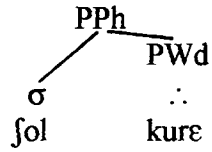
[2] Observe in (2f) that the underlying vowel /ɛ/ undergoes elision when it is preceded by a vowel. Languages vary in the strategies that they employ to correct these illicit V_1V_2 (vowel + vowel) sequences: Elision of V_1, Elision of V_2, Glide Formation or Consonant Epenthesis. Picard exhibits all four strategies to syllabify an illicit vocalic hiatus. In the cases illustrated in (02f), the preference is for Elision of V_2. For an OT approach to the subject, see Cardoso (1998b).

[3] Observe that the vowel in /ʃol/ and /dol/ is deleted when these forms are followed by a vowel-initial lexical word. Leininger (1998) proposes that this vowel/zero alternation is motivated by the highly ranked constraints INTEG-DET (the prosodic structure of the determiner should correspond with its morphological structure (Leininger 1998)), which requires that the entire clitic be in one syllable (i.e. a type of alignment), and ONSET (syllables have onsets (Prince and Smolensky 1993)).

/ dol arb /	INTEG-DET	ONSET	MAX-IO
a. ʃol . ar . mwɛr		*!	
b. ʃo . lar . mwɛr	*!		
☞ c. ʃlar . mwɛr			*

syllable (σ) (l-clitic) and the following Prosodic Word (PWd) (lexical word), within the domain of the Phonological Phrase (PPh), illustrated in (05) (irrelevant structure is omitted). I assume therefore that the analysis that I present for AWRA exclusively involves this exact prosodic configuration (see discussion in section 3.4).

(05) The Prosodic Domain of AWRA in Picard (Cardoso 1998a)



The description of the data in (02) through (04) only partly covers the complexity of AWRA. Contrary to what is implied by the forms shown in (02), AWRA does not apply categorically at the domain juncture of an l-clitic and a following consonant-initial lexical word.[4] In this context, three distinct patterns can be found: (a) faithfulness of input /l/ (/l/-preservation); (b) Across-Word Regressive Assimilation (AWRA); and (c) /l/-deletion:

(06) Variants of AWRA:

(a) /l/-preservation:	/ʃol kurɛ/	→	[ʃol **kurɛ**]	'the/this pork paté'
	/dol tart/	→	[dol **tart**]	'some pie'
(b) AWRA:	/ʃol kurɛ/	→	[ʃo**k kurɛ**]	
	/dol tart/	→	[do**t tart**]	
(c) /l/-deletion:	/ʃol kurɛ/	→	[ʃo **kurɛ**]	
	/dol tart/	→	[do **tart**]	

In the subsequent sections, I will propose a formal account for the variation aspect of the phenomenon illustrated above, within the framework of Optimality Theory.

2 Optimality Theory And Variation

Optimality Theory offers the best tools available for analyzing patterns of variation: firstly, the framework allows for multiple outputs within a single grammar through crucial unranking of constraints (cf. Reynolds 1994 and Anttila 1997); secondly, it accounts for why a certain environment favors or does not favor the application of a phonological process: since several rankings may yield the same output, quantitative values may be directly encoded in the grammar. A rule-based approach, on the other hand, does not provide the aforementioned advantages: it merely describes the phenomenon and the environments that favor the rule, without explicit quantitative predictions being present in the grammar. Besides, a rule-based approach as originally proposed by Labov (1969), does not allow for the possibility of more than two variants or outputs without the postulation of more than one rule. In the case of AWRA in Picard, two rules would be required to describe the process, i.e. /l/ → [C] (AWRA) and /l/ → ø (/l/-deletion), each followed by the respective linguistic and/or extralinguistic environments that favor the application of AWRA.

[4] Interestingly, there seems to be no variation involving the domain in which AWRA applies, i.e. the phenomenon applies exclusively within the prosodic domain indicated in this paper. An explanation for this is beyond the scope of this paper and I leave it aside for further research.

Optimality Theory is a theory of constraint interaction which advocates that a grammar consists of a set of universal constraints CON which form part of Universal Grammar. As alluded to earlier, two important premises of OT are (1) *Violability*: constraints are violable; violation of low ranked constraints occurs in order to satisfy higher ranked ones; and (2) *Ranking*: constraints are ranked on a language-particular basis; the notion of minimal violation is thus defined in terms of a language-specific ranking. Accordingly, while all constraints are present in the grammars of all languages, cross-linguistic variation can be accounted for by variation in language-specific constraint rankings. Constraints are primarily of two types: those that demand a match between the input (underlying representation, i.e. UR) and the output (surface representation) - faithfulness constraints (e.g. MAX-IO: every segment in the input has a correspondent in the output), and those that demand structurally well-formed outputs - markedness constraints (e.g. NoCoda: Codas are not allowed). In my analysis, I adopt the Correspondence Theory version of OT (McCarthy and Prince 1995), where faithfulness constraints are expressed in terms of the identity relation between input and output (in contrast to standard OT (Prince and Smolensky 1993, McCarthy and Prince 1993ab) where all constraints are stated on outputs). This way, each candidate comes from GEN (the function *generator*) with a correspondence relation that holds between the elements of the input and those of the output, and evaluation is performed in parallel on the whole candidate set. The candidate that best satisfies the constraint hierarchy of the language emerges as the optimal form.

3 AWRA And Variation

In this section, I propose an analysis to account for the variation patterns found in AWRA. As you recall from (06) in section 1, the AWRA phenomenon displays three patterns of variation: (a) /l/-preservation, (e.g. /ʃol kurɛ/ → [ʃol kurɛ] 'the pork paté'); (b) AWRA (e.g. /ʃol kurɛ/ → [ʃok kurɛ]); and (c) /l/-deletion, (e.g. /ʃol kurɛ/ → [ʃo_ kurɛ]). I begin by discussing the data collection procedures and the linguistic and extralinguistic variables that I employed in the study. This is followed by a brief discussion of the statistical program VARBRUL and its results for each variable investigated. Finally, I propose an OT analysis for the variation aspect of the AWRA phenomenon.

3.1 The data collection and the independent variables

3.1.1 The extralinguistic independent variables:

This study consists of 2,783 tokens of variants of AWRA collected in the field by Julie Auger for the Picard project during the summers of 1996 and 1997, which were further transcribed by four research assistants (including me) and later rechecked for consistency.[5] The data collected were stratified among six independent variables and later analyzed by the Varbrul program (see 3.2): three extralinguistic factor groups: (1) Level of Formality, (2) Speaker and (3) Geographic Location; and three linguistic factor groups: (1) Status of the l-Clitic, (2) Place of Articulation of the Following Consonant and (3) Manner of Articulation of the Following Consonant.

The subjects (Speakers 1-9) were nine male adult native speakers of Picard, with an average age of more than 70 years old; they inhabited five villages in the Picardie region in northern France: Feuquières, Fressenneville, Bienfay, Bouillancourt and Nibas. Women and younger speakers were not included in the investigation because the vast majority of native speakers of Picard who still use

[5] The rechecking process was necessary because at the initial stage of the transcription task, the research assistants were not aware of the AWRA phenomenon. Besides, they were following a less conservative approach in the transcription (mostly syntax-oriented), and therefore little or no attention was given to the assimilation process.

the language routinely are older men. Languages such as Picard which die out gradually via the progressive failure of intergeneration transmission, usually retreat in the final generations to a few spheres of use: they persist in domestic settings among the older (usually male) generation, and they are used for casual interaction among contemporaries who were schoolmates or coworkers in their younger years.

In order to collect tokens from a wide range of stylistic levels, the data collection methodology used in this study provides a three-level distinction in a formality hierarchy: (1) informal interview, (2) formal interview, and (3) collection of written documents.

(1) The informal interview consisted of tape-recorded conversations between the field worker and the interviewee or between the interviewee and other native speakers of Picard. To avoid more careful (or less informal) speech produced through the influence of the observer's paradox, the first ten to fifteen minutes of each conversation was ignored, and whenever possible, preference was given to data from the intercourse when only speakers of Picard were involved. Also, tokens present in discussion topics that could elicit more formal speech or code-switching (e.g. politics, history, professional and educational activities, etc.) were excluded from the analysis.

(2) The formal interview consisted of a tape-recorded translation task in which the subjects were asked to orally translate French sentences (designed for the purpose of this study) into Picard. The only way to elicit less informal oral data was through the translation task because Picard, as a dying (or recessive) language, is characterized by monostylism. As has been attested in the sociolinguistic literature (cf. Dressler 1972, Dressler and Wodak 1977, Dorian 1977, Dressler 1988, among others), languages in the process of decay are mostly used in a single formality style (either in a casual style, or in religious contexts such as church services, scripture readings, etc.). In the case of Picard, its use is limited to more casual styles, i.e. for routine topics in oral situations and for intimate interaction with close friends and at home.

(3) The collection of written documents consisted of the selection of written documents from at least one speaker from each region investigated. These documents were extracted from articles from the Picard magazine Ch'Lanchron, books (including compilations), and unpublished material (including short stories, articles and a few private letters).

The transcription of the data was not as straightforward as implied above. Because of intense bilingualism (French and Picard) in the communities investigated and high levels of code-switching (even within the same sentence!), one of the most arduous decisions I had to make was to determine whether the subject pronoun /al/ was truly the Picard form [al] 'she', rather than the French equivalent [ɛl]. Due to the phonetic similarity of the vowels [a] and [ɛ] in rapid speech (in informal interview), they often confounded me during the transcription of the data.[6] Another problem I encountered involves the determiner and pronoun /ɛl/ 'the, him/her', which also resembles the French equivalent [lə] in cases of vocalic hiatus in Picard (Cardoso 1998b). In such cases, the clitic's vowel [ɛ] is deleted and the output is similar to that of French: e.g. Picard /tu ɛl tɛ/ → [tu l tɛ̃] vs. French [tu l tɛ̃] 'all the time'. In order to remedy these problems, ambiguous occurrences of /ɛl/ in vocalic hiatus contexts and /al/ (or /ɛl/) in mostly French sentences were rejected. The final results, however, may have been affected by these limitations.

The distribution of subjects according to Age, Geographic Location, and the Level of Formality from which the data were elicited is illustrated below.

[6] It is important to mention that regressive assimilation and /l/-deletion are not found in French. Thus, confusion arises only when the full Picard form [al] is used.

(07) Table 1: Distribution of subjects according to Age, Location, and Level of Formality

Speaker	Age	Geographic Location	Level of Formality		
			Informal	Formal	Written
Speaker 1	60s	Feuquières	No	✓	✓
Speaker 2	70s	Feuquières	✓	✓	No
Speaker 3	60s	Fressenneville	✓	No	No
Speaker 4	84	Fressenneville	No	✓	✓
Speaker 5	60s	Bienfay	✓	✓	✓
Speaker 6	84	Bouillancourt	✓	✓	✓
Speaker 7	37	Nibas	✓	✓	✓
Speaker 8	80s	Nibas	✓	No	No
Speaker 9	86	Nibas	No	No	✓

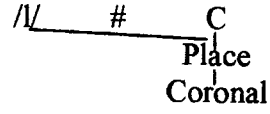
3.1.2 The linguistic independent variables

As mentioned above, three linguistic factor groups were included in the AWRA investigation: (1) Status of the l-Clitic, (2) Place of Articulation of the Following Consonant and (3) Manner of Articulation of the Following Consonant.

(1) Status of the l-Clitic factor was selected based on Debrie's (1981) description and my initial observation that the AWRA process does not apply in an identical fashion throughout the repertoire of l-clitics: while the typically Picard forms /ʃol/ and /dol/ are more likely to assimilate, the determiner and complement pronoun /ɛl/ are more likely to surface as faithful to the input, i.e. with /l/ preserved. The factors included in this group are: (a) Determiner /ʃol/, (b) Preposition /dol/, (c) Subject Pronoun /al/, (d) Preposition /al/, Complement Pronoun /ɛl/ and (e) Determiner /ɛl/.

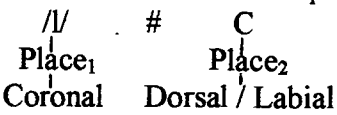
(2) Place of Articulation of the Following Consonant factor was included in this investigation based on the hypothesis that the place of articulation of the following lexical word's initial segment could have an effect on the outcome of the AWRA process. As has been well documented in the literature on syllable structure, heterosyllabic coda-onset restrictions are sensitive to place (e.g. where the coda /l/ licenses its own place vs. where place is shared with the following onset consonant). Since the clitic final consonant /l/ is coronal, it could be the case that /l/-preservation would be more likely to occur if the following consonant is also coronal (e.g. /t, s, l, n/, etc.), in which case both the coda and the onset would share the same place of articulation; (see 08a). This is in contrast to cases where the coda is followed by a non-coronal (i.e. labial (e.g. /f, v, m/, etc.) or dorsal (e.g. /k, g/)) segment. In the latter case, assimilation or /l/-deletion would be likely to occur because otherwise the coda would be required to license its own place; (see 08b).

(08) (a) /l/-preservation expected:



E.g.: /ʃol tab/ → [ʃol_tab]

(b) AWRA or /l/-deletion expected:



E.g.: /ʃol fɛt/ → [ʃof_fɛt] or [ʃo_fɛt]

Based on the discussion above, the factor group Place of Articulation of the Following Consonant includes the following factors: (a) Labial, (b) Coronal and (c) Dorsal.

(3) Manner of Articulation of the Following Consonant factor was selected based on the cross-linguistic observation that heterosyllabic coda-onset restrictions are sonority driven (i.e. onsets must be less sonorous than codas). The prediction for Picard is that the AWRA and /l/-deletion variants would be used to repair a relatively bad sonority profile (for example, while an /l-p/ coda-onset sequence may be preserved, an /l-m/ sequence may surface as [mm] or [m] because of

the bad sonority profile). In the sonority hierarchy below, the nasal /m/ is closer in sonority to the clitic-final liquid /l/ than the obstruent /p/ is.

(09) Sonority Scale (e.g. Clements 1990):
Vowels > Glides > Liquids > Nasals > Obstruents

Based on the Sonority Scale above, the factors included in the independent variable Manner of Articulation of the Following Consonant are: (a) Glide, (b) Liquid, (c) Nasal, (d) Fricative and (e) Other Obstruent.

3.2 The VARBRUL Program

This section consists of a brief introduction to the VARBRUL 2 program for DOS (Pintzuk 1988) used in the analysis of the Picard corpus. This program has been extensively used in variationist studies in linguistics because, along with GoldVarb for Macintosh computers (Rand and Sankoff 1990), it is the only one deliberately designed to handle the types of data derived from studies of language variation.[7] In Young and Bayley's (1996) terms, VARBRUL is able to manage "the distributional imbalances of linguistic features in sociolinguistic data."

The results of a VARBRUL study should be interpreted as holding over the whole of the data corpus which is being investigated and, to the extent that this is a representative sample, to all similar speakers and linguistic and extralinguistic contexts. The output of a typical VARBRUL analysis contains the following information: (1) The *raw number* (N) and the *percentage* of rule application involving each factor. These results, however, do not provide enough information since they do not express the influence of each factor independently of the others. (2) The *factor weight* measures the influence that each factor has in the process under investigation, based on the corpus analyzed. It provides the most accurate view of the likelihood of variant occurrence. It consists of a list of values associated with each factor independently of other factors in the same factor group. The value indicates the degree to which a factor promotes the occurrence of each variant for the process being investigated. The higher the value, the higher the influence of that factor in the selection of the variable output. A weight value of either 1.00 or 0.00 indicates that a given factor has a categorical influence on variation for the dependent variable investigated (e.g. the AWRA phenomenon): in the context of a group factor for which the program assigned a categorical value, a weight of 1.00 indicates that a certain variant will always occur, while a weight of 0.00 indicates that that variant will never appear. Because the AWRA phenomenon consists of three variants, the weight of .33 will be established as the watershed between the weights that enhance the likelihood of a certain variant's occurrence (above .33) and those that inhibit its appearance (below .33). (3) The *input probability* is the likelihood that each variant has of occurring in general, regardless of the specific contribution of particular factors.

The first VARBRUL run includes all the original factors as they were initially conceived based on the investigator's hypothesis. It is not uncommon, however, to find that a certain group or factor does not contribute substantially to the observed variation: e.g. near categorical results (called 'knockouts') or factor groups consisting of a single factor (called 'singletons'). Because the VARBRUL program cannot calculate the weights of factors or factor groups consisting of knockouts or singletons, it is necessary to modify the analysis, either by removing these problematic factors or factor groups, or by regrouping them with other related groups or factor groups. Also,

[7] Other probabilistic programs such as ANOVA were designed to handle data collected from controlled experimentation that result in rather balanced data.

interactive factor groups (e.g. *speaker* and *geographic location* in this study) should be analyzed independently from one another so that there is no interference of one factor over the other. The process discussed above is called recoding. In order to refine the model of variation, subsequent VARBRUL runs should be conducted until the final results contain no knockouts or singletons, and until all factors that are theoretically similar (and equally influential) are regrouped into a single factor or removed from the analysis.

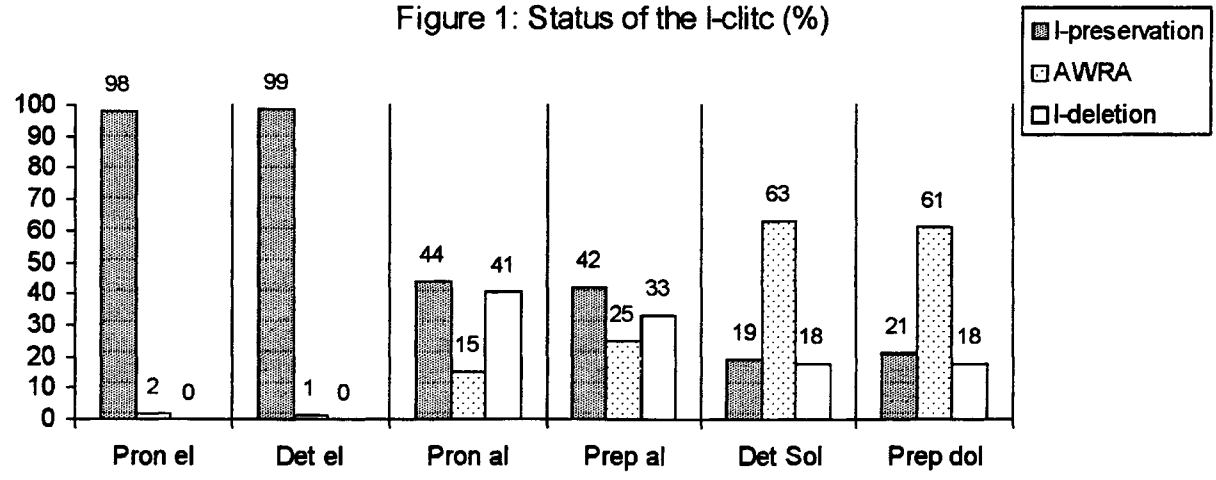
3.3 The VARBRUL quantitative analysis: results and discussion

In this section, I provide the results of the quantitative analysis conducted by the VARBRUL 2 program for the AWRA study. From all the linguistic and extralinguistic factors that I initially included in the investigation (see discussion under 3.1), VARBRUL's probabilistic results indicate that the external variables *level of formality*, *speaker* and *geographic location* and the internal variable *grammatical category of the clitic (status of the l-clitic)* have significant conditioning effects on determining the output of the AWRA phenomenon. On the other hand, the internal factors *place of articulation of the following consonant* and *manner of articulation of the following consonant* did not play a major role in the selection of the three variants, even with additional factor recodings. Because of persistent knockouts toward l-faithfulness (categorical results), the pronoun and determiner /ɛl/ were excluded from subsequent runs. Speaker 9 was also eliminated from further analyses because of his near categorical use of AWRA (98%). In the next subsections, I briefly discuss the results obtained for the linguistic and extralinguistic factors involved in the investigation of the AWRA phenomenon. An analysis of these results within the framework of Optimality Theory is provided in section 3.4.

3.3.1 The linguistic factors

The internal factor *l-clitic status* displays a pattern in which the determiner and complement pronoun /ɛl/ presented no significant variation and /l/-preservation is categorically observed. Because of their near categorical behavior, the results of the analyses in which they were included can only be given in percentages. In Figure 1, I illustrate the overall pattern detected in the first VARBRUL analysis of the six items included in this factor group.

Figure 1: Status of the l-clitc (%)



Observe the existence of three distinct patterns concerning the behavior of l-clitics: (1) as mentioned above, in the context of a complement pronoun and determiner /ɛl/ clitic, one observes near categorical /l/-preservation (98% and 99% respectively), few tokens with AWRA (2% and 1%) and categorical non-existence of /l/-deletion (0% in both cases); (2) in the context of the pronoun and preposition /al/, there is a higher percentage of /l/-preservation (44% and 42% respectively) and /l/-deletion (41% and 33%) and a lower tendency toward the AWRA variant (15% and 25%); (3) the determiner /ʃol/ as well as the determiner and preposition /dol/, on the other hand, display a pattern in which AWRA is highly observed (63% and 61% respectively), and /l/-preservation (19% and 21%) and /l/-deletion (18% and 18%) are relatively equally distributed.

Because of the categorical results obtained for /ɛl/, the clitic was excluded from further VARBRUL runs. The results of the reanalysis (3rd run)[8] for this factor group is shown in Table 2.

(10) Table 2: Grammatical status of l-clitics and AWRA (after reanalysis)

Clitic status	/l/-preservation	AWRA	/l/-deletion
Det. ʃol	.19	.58	.23
Prep. dol	.22	.55	.22
Pron. al	.45	.13	.42
Prep. al	.42	.20	.39

Observe that while the AWRA variant is favored by both the determiner /ʃol/ and preposition /dol/, the pronoun and preposition /al/ favor /l/-preservation and /l/-deletion.[9] In order to account for the three distinct patterns observed in l-clitics (see Figure 1), I acknowledge the existence of lexically specific constraints (rather than distinct constraint rankings). This will be discussed in section 3.4.1.

The numerical results achieved for the second linguistic factor *place of articulation of the following word's onset* was not consistent with my initial hypothesis, since the program did not render significant the factor group or any of the factors included. Notice in Table 3 that according to the nature of the following consonant's place of articulation, all three variants of AWRA are likely to occur.

[8] All the probabilistic results provided here derive from the final VARBRUL run without the redundant factor group *speaker* and the categorical factors *determiner and pronoun /el/*. The reason for the recodings and reanalyses will be explained as they become relevant.

[9] In her syntax-focussed investigation of subject pronouns in Picard, Auger (1993) points out that "Picard provides clear evidence of [a] case where subject pronouns have become agreement markers prefixed to finite verbs." Observe the results in percentage obtained for the grammatical status of the l-clitic in Figure 1, in which there is a relatively higher percentage of /l/-preservation for the subject pronoun clitic /al/. Based on Auger's arguments, it is possible that Picard speakers no longer recognize the subject pronoun /al/ as a clitic but instead as an affix. Recall that the data in (03) indicates that AWRA does not affect affixation. Accordingly, it is possible that Picard is moving toward a grammar in which certain proclitics are interpreted as affixes, left-aligned to Prosodic Words. Whether this constitutes an instance of language change in progress because of a conflict between the phonology and the syntax of Picard requires further investigation.

The possibility that /al/ may be an affix rather than a clitic in syntax does not pose any major problem for the analysis proposed here: the Prosodic Phonology approach adopted does not presuppose coincidence between the phonological and the syntactic components. It could be the case that in the syntax /al/ bears the status of an affix, while in the phonology it patterns with the other l-clitics due to the syntax-free prosodic status that they assume in the language. As Berendsen (1986) notes, there is a difference between prosodic cliticization and syntactic cliticization: elements can be prosodic clitics while not being syntactic ones, vice-versa, or both.

(11) Table 3: Place of articulation and AWRA

	/l/-preservation	AWRA	/l/-deletion
Labial	.38	.35	.33
Coronal	.32	.33	.35
Dorsal	.34	.34	.32

Finally, the third linguistic variable *manner of articulation of the following consonant* was selected as non-significant by the VARBRUL program, as expressed in the quantitative values in Table 4. Based on the values assigned to each factor, the manner of articulation of the following consonant does not interfere in the selection of the variants involved in AWRA, because the likelihood that each factor will interfere in the process is relatively equal. Again, my initial hypothesis that the assimilation phenomenon could be driven by sonority was not confirmed.

(12) Table 4: Manner of articulation and AWRA

	/l/-preservation	AWRA	/l/-deletion
Glide	.35	.32	.33
Liquid	.34	.35	.31
Nasal	.30	.32	.38
Fricative	.34	.33	.33
Other obstruent	.34	.35	.31

One possible explanation for the fact that segmental properties (i.e. place and manner of articulation) do not seem to affect the outcome of the AWRA process may be in the prosodization of the elements involved in the process. It could be the case that the sonority/place profile may only be relevant internal to the Prosodic Word, involving root and possibly affixes. As you recall from (05) in section 1, the AWRA phenomenon is exclusively sensitive to the Phonological Phrase domain, in the juncture of an unstressed syllable and the following Prosodic Word. Because the consonant-initial lexical word prosodicizes as a Prosodic Word, the l-C sequence is interrupted within this domain. Following the rationale developed so far and the empirical evidence from AWRA, it seems reasonable to suppose that constraints sensitive to the Prosodic Word may not have the same strength in other prosodic domains.

3.3.2 Results of the extralinguistic factors

The social dependent variable level of formality exhibited a significant effect on the AWRA phenomenon, as can be observed in Table 5.

(13) Table 5: Level of formality and AWRA

	/l/-preservation	AWRA	/l/-deletion
Written	.38	.23	.39
Formal	.41	.36	.23
Informal	.22	.40	.38

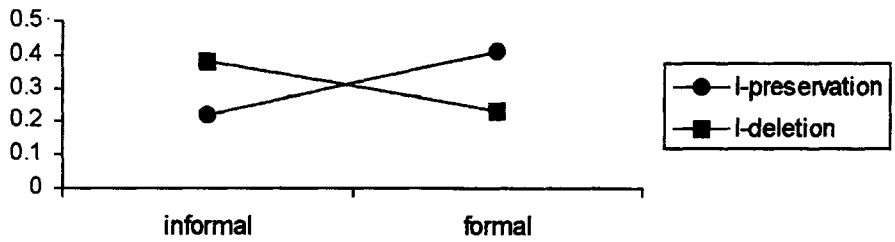
Notice that the three variants of AWRA did not behave similarly in the three stylistic levels as the differences in the trends of variation suggests: while the informal environment favors less faithful forms (i.e. AWRA and /l/-deletion), a more formal environment favors a more faithful output (i.e.

/l/-preservation and AWRA)[10]. Contrary to my expectations, the written style favors both the most faithful (i.e. /l/-preservation) and the least faithful (i.e. /l/-deletion) of all three variants, which constitutes a counterexample to van Oostendorp's (1997: 209) proposal that "[t]he more formal the style level, the higher ranked the faithfulness constraints." Also, the AWRA variant is unexpectedly favored in the formal style.

A viable explanation for the unexpected likelihood of AWRA application in more formal environments is group identity toward the AWRA variant, a characteristic marker of Vimeu Picard. Because of the current revival of Picard in the region of Vimeu and the upward progressive social status it is assuming in a community in which French prevails, some characteristically Picard forms are sometimes overused as markers of group identity. This is the case for speaker 7, for instance, who curiously produced a higher percentage of AWRA in formal situations (67% AWRA, 19% /l/-deletion and 14% /l/-preservation), and a lower percentage of AWRA in informal situations (26% AWRA, 41% /l/-deletion and 33% /l/-preservation). As for the written style, the result may reflect the fact that, unlike the oral data, the written corpus derives from written documents that span the period of approximately 30 years, and besides, it includes formal as well as informal documentation.

Focussing on the formal and informal styles separately, observe the opposite trends for /l/-preservation and /l/-deletion: while /l/-preservation increases as the contexts becomes more formal, /l/-deletion decreases in the same context; this is illustrated in Figure 2. As I will discuss in section 3.4.2.2., this suggests the postulation of different (stylistic) grammars for the AWRA phenomenon.

Figure 2: Formal and informal styles



The second extralinguistic factor *speaker* was selected as significant by the VARBRUL program. Due to the absence of /l/-preservation (knockout) in the speech of Speaker 9 (0% of /l/-preservation, 98% of AWRA and 2% of /l/-deletion), he was excluded from the second probabilistic analysis. The results involving the remaining eight speakers are illustrated in Table 6 below.

(14) Table 6: Probabilities in the speech of eight speakers

SPEAKERS	/l/-preservation	AWRA	/l/-deletion
Speaker 1	.35	.31	.34
Speaker 2	.35	.31	.34
Speaker 3	.30	.30	.40
Speaker 4	32	.28	.40
Speaker 5	.36	.26	.37
Speaker 6	.34	.30	.36
Speaker 7	.30	.45	.25
Speaker 8	.31	.48	.21

[10] For expository reasons, I consider three degrees of faithfulness: /ʃol fɛt/ → (1) [ʃol fɛt] (/l/-preservation) > (2) [ʃof fɛt] (AWRA) > (3) [ʃo fɛt] (/l/-deletion), in which the first form is more faithful to the input than the second, which is in turn more faithful than the third.

Notice that two distinct patterns can be observed: one in which all three variants are relatively equally distributed (speakers 1 through 6); and another pattern in which AWRA is favored as opposed to /l/-faithfulness and /l/-deletion (speakers 7 and 8).

As a consequence of the relevance of the social factor *speaker*, the third social variable *geographic location* should as well constitute a major effect in determining the variants of AWRA, for one factor inherently includes the other. In order to achieve a more accurate result of the factor group *geographic location*, a third analysis of the program was run without the (probably) interfering factor *speaker*. The results are illustrated in Table 7.

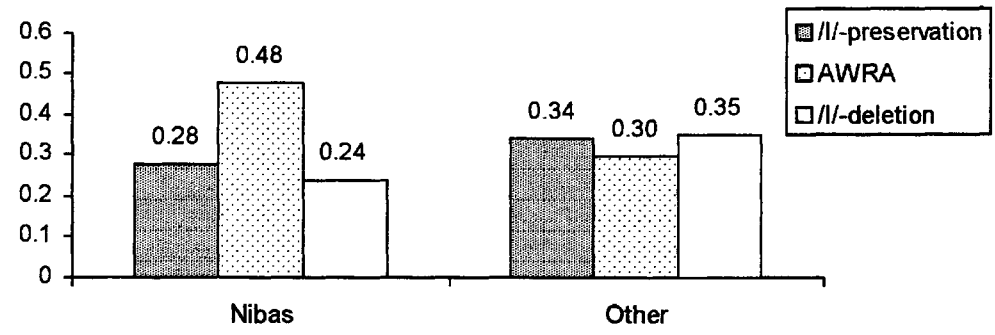
(15) Table 7: Probabilities in five geographic locations

VILLAGE	/l/-preservation	AWRA	/l/-deletion
Nibas	.28	.48	.24
Feuquières	.31	.30	.39
Fressenneville	.38	.29	.32
Bienfay	.30	.32	.38
Bouillancourt	.38	.30	.32

Observe that in the village of Nibas, the AWRA variant is more likely to appear (.48) while the two other variants are equally distributed (and equally disfavored). In the other villages, on the other hand, each variant is relatively equally expected to occur (average around .33).

Not surprisingly, this pattern conforms to the facts found in the factor group *speaker*: recall from Table 1 in section 3.1.1 that speakers 7 and 8 are from the village of Nibas while the other speakers are from different locations. For ease of exposition (and because exactly two patterns can be observed), I will regroup this factor group into two major categories: Nibas and Other (including all the remaining villages). The result (in probability) is illustrated in Figure 3.

Figure 3: AWRA and Geographic Location



Because the results for both factors *speaker* and *geographic location* indicate that the influencing factor in the variation pattern involving these two groups is *geographic location*, the variation analysis that I will present is based on the results illustrated in Figure 3 above.

3.4 The Analysis
3.4.1 Constraints and the domain-specific constraint approach

In order to account for the AWRA phenomenon, I propose an analysis within the theoretical framework of Optimality Theory, as previously discussed in section 2. In (16) below, I display the relevant constraints that I will utilize in the analysis.[11]

(16) Constraint definitions:

MAX-IO:	Every segment of S_1 has a correspondent in S_2 (McCarthy and Prince 1995)
*MC:	(No Multiple Correspondence) Elements of the input and the output must stand in a one-to-one correspondence relationship with each other (Lamontagne and Rice 1995)
NoCoda:	Codas cannot license a Root node (cf. Prince and Smolensky 1993)

MAX-IO is a family of constraints that militates against deletion. It ensures that every segment of the input (S_1) have a corresponding segment in the output (S_2). In cases of /l/-deletion, illustrated back in (06c), the optimal candidate violates MAX-IO because the input segment /l/ is not present in the output form: / ʃol vak / → [ʃo_ vak] 'the cow'.

The constraint *MC or "no multiple correspondence" in (16) requires a one-to-one relation between input and output features, thus extending correspondence to include both the segmental and featural levels, and therefore avoiding the association of input features with other Root nodes:[12]

(17) Correspondence (Lamontagne and Rice 1995)
For two Root nodes X and Y, where X is part of the input and Y the output, X and Y correspond if some features of X correspond with features of Y

For a segment of the input and output to correspond, it is necessary that they match in their feature specification for some or all of their features. In the cases of Assimilation (the AWRA variant)

---

[11] Due to limitations of space, only three constraints that directly play a role in the segmental aspects of AWRA are discussed. A relevant undominated constraint that predicts the directionality of the assimilation process is LEX-Faith, proposed by Casali (1997) and Pulleyblank (1997) (under 'Faith-Stem'), though implicit in McCarthy and Prince (1995) under the Root-Affix Faithfulness Metaconstraint: Root-Faith >> Affix-Faith. Originally, the notion that certain prominent positions maintain contrasts goes back to Trubetzkoy (1939), and has recently been discussed in the works of Steriade (1995) and Beckman (1998). LEX-Faith expresses the cross-linguistic tendency for preservation of information contained in lexical words rather than in function words. For Picard, the hierarchical position of LEX-Faith above the three constraints in (16) prevents cases of progressive assimilation (e.g. /ʃol kurɛ/ → *[ʃol lurɛ] vs. ✓[ʃok kurɛ] 'the pork paté'), as I illustrate below:

Tableau F[3]: Directionality of AWRA

/ʃol kurɛ/	LEX-Faith	MAX-IO, *MC, NoCoda
(a) ʃol . lurɛ	*!	
☞ (b) ʃol . kurɛ		*
☞ (c) ʃo . kurɛ		*
☞ (d) ʃok . kurɛ		*

[12] At least three other possibilities exist to capture the violation of multiple correspondence in cases of across-word assimilation: (1) No CrispEdge (Ito and Mester 1994): Multiple linking between prosodic categories is prohibited; (2) Align-Left-Word or "No Resyllabification" (Kiparsky 1993, Reynolds 1994): No resyllabification is allowed across word boundaries; and (3) linearity (McCarthy and Prince 1995, Pater 1996): S_1 reflects the precedence structure of S_2, and vice versa. As shown below, in cases of regressive assimilation, the precedence relation of S_1 /(l-k)/ is not reflected in S_2 [(k-k)]: /l/ precedes /k/ in S_1 but not in the output:

(i) /ʃo(l k)urɛ/ → [ʃo(k k)urɛ]

illustrated in (01) and (02) above, the input segment triggering the assimilation is in a multiple correspondence relation because it is multiply linked to two output segments: input /v/ in /ʃol vak/ → [ʃov vak], for instance, corresponds to two segments in the output and therefore violates *MC.

Finally, the NoCoda constraint expresses the cross-linguistic observation on syllabic well-formedness that coda segments are marked. In contrast to Prince and Smolensky (1993), the constraint in (16) is formulated in terms of licensing and therefore a syllable final consonant can only surface without incurring a violation of this constraint if it is licensed by a following Root node. In the cases of /l/-preservation illustrated in (06a) above, the output forms violate NoCoda because the coda /l/ bears and therefore licenses its own Root node, e.g. /ʃol vak/ → [ʃol̲ vak]. In cases of assimilation (AWRA), however, NoCoda is not violated because the assimilated coda segment is licensed by the Root node of the initial consonant of the following word, e.g. /ʃol vak/ → [ʃov vak] (coda [v] is licensed by the following onset [v]'s Root node). This is shown in the representations below, using standard Onset-rhyme theory (segments stand for Root nodes).

(18) a. Violation of NoCoda: b. Satisfaction of NoCoda:

As may have been noticed from the data provided in this paper, the three constraints above do not have the same effect in prosodic (or morphosyntactic) domains distinct from the one argued to be relevant for the AWRA phenomenon. The NoCoda constraint, for instance, does not seem to be highly ranked internal to the Prosodic Word, as I illustrate below.

(19) The NoCoda constraint at the PWd domain in Picard

(a) [bel.ʒik]$_{PWd}$ 'Belgium' (b) [mal.po.li]$_{PWd}$ 'impolite' (c) [ar.mwɛr]$_{PWd}$ 'closet'

Only between an l-clitic and the following lexical word (or at the juncture of an unstressed syllable and the following Prosodic Word, within the Phonological Phrase) does the NoCoda constraint become relevant. This behavior found with clitics in Picard conforms to a variety of studies which demonstrate that clitics behave in a peculiar way with respect to phonological processes. As Zwicky (1977) points out, clitics are "morphemes that present analytic difficulty because they are neither clearly independent words nor clearly affixes". Due to their hybrid nature, the hypothesis that there exist constraints (e.g. on syllabification) that are exclusively pertinent to this category of words is not surprising.

The claim that some constraints have a stronger effect in a larger domain than in a smaller one or vice versa is not an original claim (cf. Chomsky and Halle 1968, Selkirk 1972 et seq., Nespor and Vogel 1986, Hyman, Katamba and Walusimbi 1987, Booij 1988, Hayes 1989 and 1990, Condoravdi 1990, McCarthy and Prince 1995, Pater 1996, Buckley 1996, Pulleyblank 1997, Cardoso 1997, Peperkamp 1997, among others). Within the OT approach to prosodic phonology that I adopt in this study, there are three possibilities for dealing with such domain-specific phenomena: (1) the co-grammar or reranking approach, proposed by Itô and Mester (1995), where lexicon-internal variation is an effect of reranking of Faithfulness constraints. (2) The constraint domains approach (Buckley 1996), a variant of Itô and Mester's proposal, which presupposes the existence of different grammars (different constraint rankings) for specific sub-strings of surface representation. Applying these two views to the investigation of AWRA would require that the three constraints above and their proposed ranking hold only within the domain argued for the particular

process (i.e. at the juncture of an unstressed syllable and a Prosodic Word, within the Phonological Phrase), and other rankings (or grammars) would therefore be necessary for different prosodic domains. (3) The third approach to domain-sensitive phenomena involves domain-sensitive constraints (cf. McCarthy and Prince 1993b, Pater 1996). This constitutes a more constrained approach, as instead of the assignment of different grammars for specific prosodic domains, constraints are decomposed into distinct domain-specific constraints, each having a different hierarchical status within a single grammar.[13] I adopt this approach not only because it is more constrained but, importantly, because it harmonizes with my analysis of variation, which also postulates the existence of a single grammar to account for variable outputs within a speech community.

As a consequence of this approach, the three constraints above bear a specification for the domain in which they are relevant. For AWRA, this is the Phonological Phrase, at the domain juncture of a unstressed syllable and the following Prosodic Word. For expository reasons, I represent this domain juncture as ϕ, and each constraint will be labeled as MAX-IO$_\phi$, *MC$_\phi$ and NoCoda$_\phi$ indicating that these constraints are sensitive to the Phonological Phrase domain. In the tableaux below, I illustrate how these constraints interact with other non domain-specific constraints. Suppose that in a given Picard grammar, the hypothetical ranking NoCoda$_\phi$ >> MAX-IO >> NoCoda is responsible for the categorical selection of the outputs /l/-deletion in cliticization and /l/-preservation in prefixation.

(20) A hypothetical Picard grammar

	/ʃol fɛt/	NoCoda$_\phi$	MAX-IO	NoCoda
Cliticization: /l/-deletion	☞ (a) (ʃo_ fɛt)$_\phi$		*	*
	(b) (ʃol fɛt)$_\phi$	*!		*
	/mal poli/			
Prefixation: /l/-preservation	(c) (ma_ ʃãs)$_{PWd}$		*!	*
	☞ (d) (mal ʃãs)$_{PWd}$			*

Notice that in cliticization, the winning candidate (a) does not violate NoCoda$_\phi$ because the only coda present in this form ([t]) is outside the scope of the constraint (i.e. the domain juncture ϕ). The candidate does, however, violate the lower-ranked general NoCoda constraint due to the licensing of /t/ as a coda. A violation of the domain-sensitive NoCoda$_\phi$ is shown in the non-optimal candidate (b). In prefixation, the winning candidate (d) does not incur in violation of NoCoda$_\phi$ because the prosodization of prefix plus root sequences (i.e. ((pref)$_\sigma$ (root)$_{PWd}$)$_{PWd}$ apud Cardoso 1998a) does not constitute the domain in which this domain-sensitive constraint operates. Notice, however, that candidate (d) incurs in two violations of the general NoCoda constraint because of the licensing of the two codas [l] and [s]. Nevertheless, it is selected as optimal over candidate (c) because deletion of coda [l] in (c) violates the more highly ranked constraint MAX-IO.

For expository convenience, I will only refer to the distinction between the general constraints (e.g. NoCoda) and the domain-specific constraints (e.g. NoCoda$_\phi$) when they become relevant to the analysis.

[13] As Pater (1996) points out, "the tack of proliferating constraints over that of proliferating grammars [is preferable] because [...] it gives a clearer view of the limits that a language imposes on reranking, and especially because the proliferation of lexically specific constraints seems independently necessary (e.g. alignment constraints [...])."

3.4.2 The categorical results

In this section, I provide an analysis for the categorical results obtained for the clitic /ɛl/ and for Speaker 9. As you recall from the previous section, the clitic (determiner and pronoun) /ɛl/ presented a near categorical result in which the variant /l/-preservation occurred 98.5% of the times, while Speaker 9 demonstrated a categorical tendency toward the AWRA variant (98%).

3.4.2.1 The determiner and pronoun /ɛl/

The categorical results obtained for the clitic /ɛl/ (Group [ɛl] henceforth) (98.5% of /l/-preservation, 0% of /l/-deletion and 1.5% of AWRA) may suggest one or both of the following: (1) due to the striking similarity between the Picard form /ɛl/ and the French equivalent /lə/ (especially in cases of vocalic hiatus – see discussion in 3.1.1), speakers code-switch to the French grammar in which AWRA and /l/-deletion do not occur, and the French form surfaces; or (2) speakers do not formally represent the clitic /ɛl/ as containing an underlying vowel /ɛ/. Julie Auger (personal communication) pointed out the possibility that the underlying form for such clitics could be /l/, and [ɛ] would be epenthesized in order to correct an unsyllabifiable consonant cluster (e.g. /pur l fwɛr/ → [pur ɛ̲l fwɛr] vs. *[pur.lfwɛr], *[purl.fwɛr] 'to do it').

The possibility that the clitic /ɛl/ is underlyingly a single segment /l/ can account for why /l/-deletion never appears in the results. If /l/ is the only segment, its deletion will leave no remaining segmental trace of the clitic. Its presence, therefore, must be forced by a functional constraint that requires that monosegmental morphemes be preserved in the output. The constraint MAX-MS, proposed by Schuh (1996) and Casali (1997), requires the preservation of features in monosegmental morphemes. Assuming that this constraint is highly ranked in the grammar of Picard, we can account for the absence of /l/-deletion when the clitic /ɛl/ is involved.

(21) MAX-MS:
Every input segment which is the only segment in its morpheme must have a corresponding segment in the output (Schuh 1996, Casali 1997).

Because /ɛl/ clitics do not undergo assimilation or deletion, it is reasonable to assume that they are lexically marked for /l/-preservation.[14] In order to restrict the application of regressive assimilation within this group of clitics, and still preserve the hypothesis that we are dealing with a single grammar, a lexically specific version of the constraint *MC is required: $*MC_{[el]}$, where subscripted [el] represents the constraint specification for the group of /ɛl/ clitics. The crucially-ranked hierarchy in (22) yields the expected results.[15]

(22) Lexically marked Group [el]: MAX-MS, $*MC_{[el]}$ >> NoCoda, MAX-IO, *MC

[14] /l/-preservation constitutes a type of "no application" (McCarthy and Prince 1995). According to the authors, for a markedness constraint (e.g. NoCoda) to be active in a language, it must dominate some relevant constraint on I-O faithfulness (e.g. MAX-IO, *MC). In the ranking provided in (22), notice that the markedness constraint NoCoda is ranked lower than the faithfulness constraints MAX-MS and $*MC_{[el]}$. This results in /l/-preservation, or no application of /l/-deletion or AWRA.

[15] In the tableau, the parenthesized vowel represents the epenthetical vowel inserted to correct the unsyllabifiable consonant cluster. The formalization of epenthesis in Picard is beyond the scope of this paper and will not be addressed.

(23) Tableau 2: Monosegmental /l/ and /l/-preservation

/l fis/	MAX-MS	*MC$_{[el]}$	NoCoda	MAX-IO	*MC
(a) (ɛ).fis	*!		*	*	
☞ (b) (ɛ)l. fis			**		
(c) (ɛ)f.fis		*!	*		*

Observe that the selected candidate is the one in which there is complete faithfulness between input and output. Candidate (b) wins because it only minimally violates the NoCoda constraint (due to the licensing of the /l/ and /s/ as codas), lowly ranked in the constraint hierarchy proposed. The remaining candidates each violate a highly ranked constraint and therefore are ruled out as optimal forms: candidate (a) violates MAX-MS because the only segment in the input is deleted in the surface form, while candidate (c) violates *MC$_{[el]}$ due to the multiple correspondence of the segment /f/ in the output.

If the underlying form for such clitics is the one in which the vowel /ɛ/ is part of the input, the rationale is identical to the discussion above. MAX-MS, however, is replaced by MAX-IO$_{[el]}$, lexically marking the clitic not to undergo deletion. The result is shown below.

(24) Tableau 3: AWRA and clitic /ɛl/ - Categorical /l/-faithfulness

/ɛl fix/	MAX-IO$_{[el]}$	*MC$_{[el]}$	NoCoda	MAX-IO	*MC
(a) ɛ . fis	*!		*	*	
☞ (b) ɛl . fis			**		
(c) ɛf . fis		*!	*		*

3.4.2.2 Speaker 9

The categorical results observed for Speaker 9 (98% of AWRA, 2% of /l/-deletion and 0% of /l/-preservation) suggest that we are dealing with a categorical grammar in which only one form results: AWRA. One could argue that these results have been affected by the fact that only written data were collected from this speaker. A closer look at the overall results, however, suggests that there is something particular about Speaker 9, since no other speaker presents such a great discrepancy between the three output forms found in written material (this can be seen in the results of the stylistic level *written* in which the distribution of each variant's occurrence is relatively similar: 36% of /l/-preservation, 32% of AWRA and 32% of /l/-deletion). To account for the nearly categorical results observed for Speaker 9, I propose the constraint ranking in (25) which results in the selection of the AWRA candidate as optimal, shown in Tableau 4.

(25) Speaker 9's categorical grammar: NoCoda$_{\phi}$, MAX-IO$_{\phi}$ >> *MC$_{\phi}$

(26) Tableau 4: AWRA in Speaker 9's grammar

/ʃol kurɛ/	NoCoda$_{\phi}$	MAX-IO$_{\phi}$	*MC$_{\phi}$
(a) ʃol . kurɛ	*!		
(b) ʃo . kurɛ		*!	
☞ (c) ʃok . kurɛ			*

Candidates (a) and (b) each violate a highly ranked constraint in Speaker 9's grammar, and therefore, both are ruled out as optimal forms: Candidate (a) violates NoCoda due to the licensing of the consonant /l/ as a coda; candidate (b), on the other hand, violates MAX-IO since input /l/ is absent from the surface form. Candidate (c) represents the optimal form found in Speaker 9's

speech, even though it violates the lowly ranked constraint *MC. Inasmuch as optimal candidates incur minimal violation marks in OT, candidate (c) is the selected output and corresponds to the facts found in speaker 9's data.

In the following section, I will propose an OT analysis for the variation pattern observed in the AWRA phenomenon using the same set of constraints introduced in (16). As I will demonstrate, the only distinction between categorical and variable phenomena lies in the demand that a variation grammar imposes on ranking, namely the crucial unranking of constraints.

3.4.3 The results of variation

In this section, I introduce the topic of how variation can be encoded in Optimality Theory and then provide an analysis for the significant results that involve variation in AWRA. As you will recall from the previous section, the internal variable *grammatical category of the l-clitic (status of the l-clitic)* and the external variables *level of formality*, *geographic location* and *speaker* all have significant effects on determining the output of the AWRA phenomenon.

3.4.3.1 Variation in Optimality Theory

As I have demonstrated in section 2, the framework of Optimality Theory provides the best tools for analyzing variation because: (a) it allows for multiple outputs within a single grammar through crucial unranking of constraints, without the need to resort to separate rules for each distinct output; (b) it expresses how a certain environment favors or does not favor the application of a phonological process; and (c) it allows for quantitative values to be directly encoded in (and therefore predicted by) the grammar. In order to achieve these goals, two different proposals have been made: (1) Anttila's (1997), a more constrained approach in which variation grammars constitute partial orders, i.e. some subset of constraints S' within the set S are all crucially unranked with respect to each other (e.g. $\{A >> \{B; C; D\}_{S'} >> E\}_S$); and (2) Reynolds' (1994), for whom variation grammars are formed of (variably ranked) floating constraints, where within the set S, some subset of constraints S' may be variably ranked with respect to some other subset S". Within each subset, constraints may be crucially unranked (i.e., floating with respect to each other) (e.g. $\{A >> \{\{B\}_{S'}; \{C; D\}_{S''}\} >> E\}_S$, in which $\{B\}_{S'}$ floats with respect to the subset $\{C; D\}_{S''}$, whose members C and D are crucially unranked with respect to each other). In terms of the options that are possible in each of the approaches above, Anttila's approach constitutes a subset of Reynolds' floating constraint approach.

Based on empirical evidence from Picard, I adopt Reynolds' approach to the subject. Anttila (personal communication) suggests that the results found in Picard are "(i) either a counterexample to his claim that [variation] grammars are partial orders, or (ii) a case of dialect mixture."

In Reynolds' (1994) conception of variably ranked Floating Constraints (FCs) and consequent encoding of probability predictions in the grammar, a constraint or set of constraints may fall anywhere within a range (set by the grammar) in the ranking hierarchy of a single language.

(27) Reynolds' (1994) variably ranked Floating Constraints (slightly modified):
A particular constraint X may be classified as being ranked somewhere within a certain range lying between two other constraints W and Z, without the specification of its exact ranking relative to a certain other constraint Y (or constraints Y_1, Y_2, etc.).
ConW >> {{ ConX } { ConY$_1$, ConY$_2$, ConY$_n$ } } >> ConZ

From the number of rankings allowed by a set of variably ranked constraints, distinct outputs can be predicted. Anttila (1997) demonstrates that the probability of each variant's occurrence is the result of the total number of rankings (or tableaux) generated by the variably ranked constraints, divided by the number of rankings for which each variant wins.

(28) Variant Probabilistic Prediction (Anttila 1997):
  (a) A candidate is predicted by the grammar iff it wins in some tableaux.
  (b) If a candidate wins in *n* tableaux and *t* is the total number of tableaux, then the candidate's probability of occurrence is *n/t*.

To illustrate, suppose that in a given grammar, GRAM, two constraints B and C float with respect to each other. This is indicated by the semi-colon (to distinguish crucial non-ranking from cases of indeterminate ranking) between the two constraints involved, with the curly brackets delimiting the set of floating constraints. As a result, the two different constraint rankings in (29b) are possible.

(29) A variably ranked grammar:
  (a) Constraint ranking: A >> {B; C} >> D
  (b) Possibilities of rankings: (a) A >> **B** >> **C** >> D
  (b) A >> **C** >> **B** >> D

Imagine that two optimal forms are possible in GRAM, i.e. $Cand_1$ and $Cand_2$. $Cand_1$ is selected when B is ranked higher than C, while $Cand_2$ is selected in the reverse situation. This is illustrated in the two tableaux in (30).

(30) Tableau (a) A >> B >> C >> D

	A	B	C	D
☞ $Cand_1$			*	
$Cand_2$		*!		

Tableau (b) A >> C >> B >> D

	A	C	B	D
$Cand_1$		*!		
☞ $Cand_2$			*	

Following Anttila's (1997) Variant Probabilistic Prediction, the variable ranking of constraints B and C results in a pattern in which two outputs are possible, and the probability of each output occurrence can be predicted by (28). For example, candidates 1 and 2 in (30) win in exactly one tableau each (n=1), and two is the total number of tableaux (t=2). n/t = 1/2 = 0.5 or 50%. Each candidate's probability of occurrence is thus 0.5 and each variant is likely to occur 50% of the time in the same grammar.

3.4.3.2 The grammar of variation

(a) Grammatical category of the /l/-clitic

The only linguistic factor group selected as significant by the VARBRUL program was the grammatical category of the /l/-clitic. According to the results, the six /l/-clitics included in this investigation can be grouped into three classes according to their behavior toward the AWRA phenomenon: (1) those in which /l/-preservation is categorically expected (discussed in the previous section), i.e. the determiner and complement pronoun /ɛl/; (2) those in which AWRA is highly favored (average of .56) and the two other variants are equally less favored (average of .21 for /l/-preservation and .23 for /l/-deletion), i.e. the determiner /ʃol/ and the preposition /dol/; and (c) those in which /l/-preservation and /l/-deletion are equally highly favored (average of .43 for /l/-preservation and .41 for /l/-deletion), while AWRA is less likely to occur (average of .17), i.e. the pronoun and preposition /al/.

In order to account for the two contrasting sets of results involving the clitics /ʃol/ and /dol/ (Group [ol] henceforth) and the two /al/ clitics (Group [al] henceforth), I propose the constraint rankings in (31) below. To account for the distinct behavior found in these two groups of clitics, we must stipulate that each group is lexically marked for a specific pattern: while Group [ol] is lexically marked to display a higher likelihood of AWRA as opposed to the other two variants, Group [al] is lexically marked to exhibit a pattern in which the variants /l/-preservation and /l/-deletion are highly and equally favored as opposed to the variant AWRA (this is illustrated by a subscripted [ol] or [al] respectively).

(31) (a) Lexically marked Group [ol]: {{ MAX-IO$_{[ol]}$; NoCoda$_{[ol]}$}; *MC$_{[ol]}$}
(b) Lexically marked Group [al]: { MAX-IO$_{[al]}$; NoCoda$_{[al]}$; *MC$_{[al]}$; *MC }

Following Reynolds' (1994) conception of variably ranked constraints, the distinct behavior observed in Groups [ol] and [al] can be accounted for if we assume that [ol] clitics result from the constraint ranking in (31a) above, in which the constraints MAX-IO$_{[ol]}$ and NoCoda$_{[ol]}$ float in relation to one another, and form a subset of constraints which floats in relation to *MC$_{[ol]}$. In the constraint ranking responsible for the output observed in Group [al] (31b), on the other hand, four constraints float with respect to each other.

Applying Anttila's (1997) variant probability prediction in (28), the results illustrated in Table 8 are obtained. Observe that under each variant, the left column indicates the number of rankings (or tableaux)[16] for each clitic Group (i.e. [ol] or [al]) in which that candidate is the winner, and the right column indicates the probability of each variant's occurrence, calculated by the formula *n/t* in (28). Under each case, notice the actual probability (averaged from Table 2 in (10)) observed for each variant in its respective group.

(32) Table 8: Prediction and actual probability of Variant Occurrence (Clitic Status)

Clitic status	Total number of tableaux	Number of tableaux / Prediction: *n/t*					
		[l]-faithfulness		AWRA		[l]-deletion	
Group [ol]	04	01	.25	02	.50	01	.25
	Actual probability:	.21		.56		.23	
Group [al]	24	10	.42	04	.17	10	.42
	Actual probability:	.43		.17		.41	

For illustrative purposes, I will demonstrate how the ranking responsible for the results in Group [ol] (31a) determines the selection of each of the three variants involved in the AWRA phenomenon, and predicts the probability of each variant to occur. According to Table 8, the two subsets of floating constraints in {{ MAX-IO$_{[ol]}$, NoCoda$_{[ol]}$}, *MC$_{[ol]}$} yield 4 rankings. From these, one yields the /l/-faithful candidate (*t/n* → 1/4 = .25), two rankings result in the AWRA candidate (*t/n* → 2/4 = .50) and the remaining ranking results in /l/-deletion (*t/n* → 1/4 = .25). Below, I illustrate the corresponding tableaux for each variant selected by the variable ranking established for Group [ol].

[16] The tableaux are merely illustrative devices. What is crucial in the analysis is the number of possible rankings established by the variably ranked constraint set.

(33) Output selection for Group [ol]: {{ MAX-IO$_{[ol]}$; NoCoda$_{[ol]}$}; *MC$_{[ol]}$}

A. /l/-preservation *(NoCoda dominated):* (1) *MC$_{[ol]}$ >> MAX-IO$_{[ol]}$ >> NoCoda$_{[ol]}$

B. AWRA *(*MC dominated):* (1) NoCoda$_{[ol]}$ >> MAX-IO$_{[ol]}$ >> *MC$_{[ol]}$
(2) MAX-IO$_{[ol]}$ >> NoCoda$_{[ol]}$ >> *MC$_{[ol]}$

C. /l/-deletion *(Max-IO dominated):* (1) *MC$_{[ol]}$ >> NoCoda$_{[ol]}$ >> MAX-IO$_{[ol]}$

(34) Tableau 9: /l/-preservation, ranking (1) in 52.A

/ʃol kurɛ/	*MC$_{[ol]}$	MAX-IO$_{[ol]}$	NoCoda$_{[ol]}$
☞ (a) ʃol . kurɛ			*
(b) ʃo . kurɛ		*!	
(c) ʃok . kurɛ	*!		

(35) Tableau 10: AWRA, rankings (1) and (2) in 52.B

/ʃol kurɛ/	NoCoda$_{[ol]}$	MAX-IO$_{[ol]}$	*MC$_{[ol]}$
(a) ʃol . kurɛ	*!		
(b) ʃo . kurɛ		*!	
☞ (c) ʃok . kurɛ			*

	MAX-IO$_{[ol]}$	NoCoda	*MC
(a) ʃol . kurɛ		*!	
(b) ʃo . kurɛ	*!		
☞ (c) ʃok . kurɛ			*

(36) Tableau 11: /l/-deletion, ranking (1) in 52.C

/ʃol kurɛ/	*MC$_{[ol]}$	NoCoda$_{[ol]}$	MAX-IO$_{[ol]}$
(a) ʃol . kurɛ		*!	
☞ (b) ʃo . kurɛ			*
(c) ʃok . kurɛ	*!		

In sum, the factor group *grammatical category of the /l/-clitic* indicates the existence of a two-fold variation pattern: one in which AWRA is highly favored while the remaining variants are equally less likely to occur (i.e. /ʃol/ and dol/), and one in which the variants /l/-preservation and /l/-deletion are more likely to occur as opposed to the AWRA variant. Through the variable ranking of the relevant constraints responsible for the output in each group of clitics, I was able to account for the different patterns found involving the groups of clitics [ol] and [al], as well as establish the probability of each variant's occurrence in its respective clitic group.

(b) Level of formality

The second significant factor group selected by the VARBRUL program was *level of formality*. The probabilistic results indicate a pattern in which the AWRA and /l/-deletion variants are favored in more informal environments, while the variants /l/-preservation and AWRA are favored as the context becomes more formal. Surprisingly, the written stylistic level demonstrates that both the most faithful /l/-preservation and the least faithful /l/-deletion variants are favored as opposed to the AWRA variant. In order to account for these variation patterns, I propose the constraint rankings in (37) for each of the stylistic levels investigated.

(37) Level of formality and AWRA

(a). Informal style: { MAX-IO$_\phi$; NoCoda$_\phi$; *MC$_\phi$; NoCoda }

(b). Formal style: { MAX-IO$_\phi$; NoCoda$_\phi$; *MC$_\phi$; MAX-IO }

(c). Written style: { MAX-IO$_\phi$; NoCoda$_\phi$; *MC$_\phi$; *MC }

Considering that speakers are more concerned with the listener's perception in more formal situations, it is reasonable to assume that faithfulness constraints (e.g. MAX-IO, *MC) predominate in such contexts for the convenience of the addressee (cf. Oostendorp 1997, Taler 1997, among others); in less formal situations, however, the fact that they become less prominent is not surprising. The results achieved for the written stylistic level contradict these expectations as it favors both the most faithful and the least faithful forms (see discussion in 3.3.2).

Observe in (37) above that, contrary to the previous analysis, three distinct grammars are necessary to account for the variation patterns found involving level of formality. As you recall from Figure 2 in section 3.3.2, the patterns observed for informal and formal levels suggests the existence of two distinct grammars: one in which /l/-preservation is favored (formal), and one in which /l/-deletion is favored (informal). Each stylistic level should therefore constitute a discrete grammar to which Picard speakers code-switch according to the context of the intercourse. This claim is in accordance with the commonly accepted assumption that every style level is a separate grammar and, consequently, a language system consists of several style-level grammars (cf. Selkirk 1972, Oostendorp 1997).

The application of Anttila's variant probability prediction (i.e. *n/t*) in (28) results in the numerical values illustrated in Table 9 below. Observe that the values provided correspond to the ones predicted by the ranking in each respective stylistic level.

(38) Table 9: Prediction and actual probability of Variant Occurrence (Level of Formality)

Level of formality	Total number of tableaux	Number of tableaux / Prediction: *n/t*					
		[l]-faithfulness		AWRA		[l]-deletion	
Informal grammar	24	04	.17	10	.42	10	.42
	Actual Probability:	.22		.40		.38	
Formal grammar	24	10	.42	10	.42	04	.17
	Actual Probability:	.41		.36		.23	
Written grammar	24	10	.42	04	.17	10	.42
	Actual Probability:	.38		.23		.39	

(c) Geographic location

The last significant factor group in the AWRA investigation is *geographic location*. As discussed in section 3.3.2, two distinct patterns could easily be delineated: while in the region of Nibas the AWRA variant is more likely to appear while the two other variants are equally distributed (and equally disfavored), in the other villages the three variants are equally expected to occur. To account for the disparity of results observed involving the factor *geographic location*, I propose the two distinct grammars below, composed of ϕ-specific constraints (see discussion in 3.4.1): one for the village of Nibas and one for the other villages (excluding Nibas). The application of Anttila's variant probability prediction yields the results illustrated in Table 10.

(39) Geographic Location and AWRA

a. Nibas (speakers 7-8): {{MAX-IO$_\phi$; NoCoda$_\phi$}; *MC$_\phi$}

b. Other villages (speakers 1-6): {MAX-IO$_\phi$; NoCoda$_\phi$; *MC$_\phi$}

(40) Table 10: Prediction & actual probability of Variant Occurrence (Geographic Location)

Geographic location	Total number of tableaux	Number of tableaux / Prediction: *n/t*					
		[l]-faithfulness		AWRA		[l]-deletion	
Nibas grammar	04	01	.25	02	.50	01	.25
	Actual Probability:	.28		.48		.24	
Other grammar	06	02	.33	02	.33	02	.33
	Actual Probability:	.34		.30		.35	

The variable ranking of the constraints in the grammar of the other villages results in a pattern in which each of the three variants of the AWRA phenomenon is equally expected to surface (probability.33). In the village of Nibas, on the other hand, the variable ranking of the constraints predicts a variation pattern in which the AWRA variant is more often favored in relation to the other variants. As shown in Table 10, the predictions made here correspond to the observed results.

4 Concluding Remarks

In this paper, I have attempted to demonstrate how Optimality Theory can serve as a framework for analyzing variation: it not only allows the possibility of multiple outputs, but also allows the possibility of predictability of occurrence for each variant involved in the variation process. The claim that the probability of each variant's occurrence may be encoded in (and therefore predicted by) the grammar yields important consequences for the study of variation and linguistic theory in general, because it constitutes an attempt to narrow down the distinction between competence and performance.

In traditional (non-variationist) linguistics, the focus of linguistic theory lies almost entirely on the "ideal speaker-listener, in a completely homogeneous speech community" (Chomsky 1965), and data collection procedures rely almost exclusively on grammaticality judgements given by individual speakers which may sometimes be biased toward more prescriptive forms. In variationist linguistics, on the other hand, the focus has been on the analysis of large corpora of spontaneous or controlled discourse, in which "errors" can be (and sometimes are) classified as systematized ways of saying the same thing. In order to achieve the correct description and/or explanation of linguistic data, systematic quantitative generalizations about language should be included in what is usually referred to as competence. In agreement with Frisch (Variation in OT: discussion group on the internet, 1997), we should be seeking the correct explanations regardless of their quantitative or categorical character – "the facts determine what the grammar looks like."

By proposing an analysis in which variables as well as the predictability of each variant's occurrence are encoded in the grammar, and therefore into competence, we obtain a more accurate and comprehensive approach to the study of language. My analysis (among many others in the sociolinguistic literature) presupposes that variation is an inherent part of what is normally referred to as competence. As a consequence, the competence that I strive to account for in this study includes much more than what Chomsky (1965) proposes to be competence. As Labov (1972: 226) points out,

> [t]he ability of human beings to accept, preserve, and interpret rules with variable constraints is clearly an important aspect of their linguistic competence or *langue*. But no one is aware of this competence, and there are no intuitive judgements accessible to reveal it to us. Instead, naive perception of our own and others'

behavior is usually categorical, and only careful study of language in use will demonstrate the existence of this capacity to operate with variable rules.

In this paper, I have provided an account for the variation patterns found for the AWRA process involving the significant factors */l/-clitic status, level of formality* and *geographic location*. In order to account for the variable results involving these factors, I have proposed the decomposition of general constraints into lexically-marked and domain-sensitive constraints, each prominent at the relevant prosodic domain of AWRA application. I have argued that variation in AWRA involves the existence of two variable grammars: one that equally favors the three variants found in the process, and one that favors AWRA over the remaining variants. I have also shown that the distinction between Speaker 9's categorical grammar and that of the other speakers lies in the limitations imposed by his fixed-ranked grammar whose output is a single candidate - AWRA. Based on predictions determined by the variable ranking of the relevant constraints in the grammar of Picard, I was able to quantitatively establish the probability of application of each variant in both grammars.

REFERENCES

ANTTILA, ARTO. 1997. Deriving variation from grammar: a study of Finnish genitives. In F. Hinskens, R. van Hout and L. Wetzels (eds.), *Variation, Change and Phonological Theory*, John Benjamins, pp.35-68.

AUGER, JULIE. 1993. On the nature of subject clitics in Picard. *Issues and Theory in Romance Linguistics*, pp.159-179.

BECKMAN, JILL. 1998. *Positional Faithfulness.* Ph.D. dissertation, University of Massachusetts Amherst.

BERENDSEN, EGON. 1986. *The Phonology of Cliticization.* Foris Publications, Dordrecht, Holland.

BOOIJ, GEERT. 1988. Review of Nespor and Vogel (1986). *Journal of Linguistics 24,* pp.515-525.

BUCKLEY, EUGENE. 1996. Levels vs. domains: the case of Kashaya vowel length. In *Proceedings of the Berkeley Linguistic Society.* California.

CARDOSO, WALCIR. 1998a. The prosodic domain of across-word regressive assimilation in Picard: an optimality theoretic account. In Jonz, J. (ed.), *Southwest Journal of Linguistics,* Vol. 17(2).

CARDOSO, WALCIR. 1998b. Vocalic hiatus resolution in Picard. Paper presented at SCIL 10, Northwest University, Evanston, Illinois, U.S.A. To be published in the MIT Working Papers in Linguistics - Proceedings of SCIL 10.

CASALI, RODERIC. 1996. *Resolving Hiatus.* Ph.D. dissertation, UCLA.

CHOMSKY, NOAM. 1965. *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT Press.

CHOMSKY, NOAM & MARK HALLE. 1968. *The Sound Pattern of English.* New York: Harper and Row.

CLEMENTS, G. 1990. The role of the sonority cycle in core syllabification. In J. Kingston and M. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*,. Cambridge University Press, pp.283-333.

CONDORAVDI, CLEO. 1990. Sandhi rules of Greek and prosodic theory. In S. Inkelas and D. Zec (eds.), *The Phonology-Syntax Connection.* Chicago: The University of Chicago Press.

DEBRIE, R. 1981. Problèmes posés par la présence de l'assimilation régressive dans le sud-ouest du domaine picard. *Revue de Linguistic Romane,* vol. 45-2, pp.422-464.

DORIAN, N. C. 1977 The problem of the semi-speaker in language death. *International Journal of the Sociology of Language 12:* 23-32.

DRESSLER, WOLFGANG. 1972. On the phonology of language death. *Papers of the Chicago Linguistic Society 8*: 448-57.

DRESSLER, WOLFGANG & R. WODAK-LEODOLTER. 1977. Language death. *Journal of the Sociology of Language 12:* 33-51.

DRESSLER, WOLFGANG. 1988. Language death. In Frederick J. Newmeyer, ed., *Linguistics: The Cambridge Survey.* Vol. IV. Language: The Socio-Cultural Context. Cambridge University Press. Pp. 184-92.

FRISCH, STEFAN. 1997. [Variation in OT: discussion group on the internet]. Indiana University.

GUY, GREGORY. 1997. The grammar of variation. In F. Hinskens, R. van Hout and W. Wetzels (eds.), *Variation, Change and Phonological Theory*, John Benjamins, pp.125-143.

HAYES, BRUCE. 1989. The prosodic hierarchy in meter. In P. Kiparsky and G. Youmans (eds.), *Rhythm and Meter.* Orlando: Academic Press.

HAYES, BRUCE. 1990. Precompiled phrasal phonology. In S. Inkelas and D. Zec (eds.), *The Phonology-Syntax Connection.* Chicago: The University of Chicago Press.

HYMAN LARRY, FRANCIS KATAMBA & L. WALUSIMBI. 1987. Luganda and the strict layer hypothesis. *Phonology Yearbook 4:* pp.87-108.

ITÔ, JUNKO & RALF-ARMIN MESTER. 1994. Realignment. In Proceedings of the June 1994 Utrecht Prosodic Morphology Workshop.

ITÔ, JUNKO & RALF-ARMIN MESTER. 1995. The core-periphery structure of the Lexicon and Constraints on reranking. In *Papers in Optimality Theory.* UMass, GLSA, pp.181-209.

KIPARSKY, PAUL. 1993. Variable rules. Paper presented at the Rutgers Optimality Workshop, New Brunswick, N.J.

LABOV, WILLIAM. 1969. Contraction, deletion, and inherent variability of the English copula. *Language,* 45: 715-762.

LABOV, WILLIAM. 1972. *Sociolinguistic Patterns.* Philadelphia: University of Pennsylvania Press.

LAMONTAGNE, GREG & KEREN RICE. 1995. A correspondence account of coalescence. *Papers in Optimality Theory.* UMass, GLSA, pp.211-223.

LEININGER, ALISON. 1998. *Regressive Assimilation in Vimeu Picard: an OT account.* Ms. Indiana University.

MCCARTHY, JOHN & ALAN PRINCE. 1993a. *Prosodic Morphology I. Constraint Interaction and Satisfaction.* Ms. University of Massachusetts, Amherst, and Rutgers University.

MCCARTHY, JOHN & ALAN PRINCE. 1993b. Generalized Alignment. In Geert Booij and Jap van Marle (eds.), *Yearbook of Morphology 1993,* 79-153. Dordrecht: Kluwer Academic Publishers.

MCCARTHY JOHN & ALAN PRINCE. 1995. Faithfulness and reduplicative identity. *Papers in Optimality Theory.* UMass, GLSA, pp.249-384.

NAGY, NAOMI & WILLIAM REYNOLDS. 1997. Optimality Theory and variable word-final deletion in Faetar. *Language Variation and Change 9,* pp.37-55.

NESPOR, MARINA & IRENE VOGEL. 1986. *Prosodic Phonology.* Dordrecht: Foris.

OOSTENDORP, MARC VAN. 1997. Style levels in conflict resolution. In F. Hinskens, R. van Hout and W. Wetzels (eds.), *Variation, Change and Phonological Theory,* John Benjamins, pp.207-229.

PATER, JOSEPH. 1996. *Consequences of Constrain Ranking.* Ph.D. dissertation, McGill University.

PEPERKAMP, SHARON. 1997. *Prosodic words.* HIL dissertation 34, The Hague: Holland Academic Graphics.

PINTZUK, SUSAN. 1988. *VARBRUL programs* [computer program]. Philadelphia: University of Pennsylvania Department of Linguistics.

PRINCE, ALAN & PAUL SMOLENSKY. 1993. *Optimality theory: constraint interaction in generative grammar.* MIT Press, Cambridge, Mass.

PULLEYBLANK, DOUGLAS. 1997. Optimality theory and features. In *Optimality Theory - An Overview,* D. Archangeli and D. T. Langedoen (eds.), Blackwell Publishers.

RAND, DAVID & DAVID SANKOFF. 1990. *GoldVarb: a variable rule application for the Macintosh (version 2.0)* [Computer program]. Montréal: Centre de recherches mathématiques, Université de Montréal.

REYNOLDS, WILLIAM. 1994. *Variation and phonological theory.* Ph.D. thesis, University of Pennsylvania.

SAPIR, EDWARD. 1930. Southern Paiute, a Shoshonean language. *Proceedings of the American Academy of Arts and Sciences* 65.1.

SELKIRK, ELISABETH O. 1972. *The Phrase Phonology of English and French.* Ph.D. thesis, MIT. Distributed in 1981 by the Indiana University Linguistics Club, Bloomington, Indiana.

SCHUH, RUSSELL. 1995. Aspects of Avatime phonology. In *Studies in African Linguistics 24.* pp.31-67.

STERIADE, DONCA. 1995. Underspecification and markedness. In J. Goldsmith (ed.), *The Handbook of Phonological Theory,* Cambridge, Mass.: Blackwell Publishers.

TALER, VANESSA. 1997. *S-Weakening in the Spanish of San Miguel, El Salvador.* MA thesis, McGill University.

TRUBETZKOY, N. S. 1939. *Grundzuge der phonologie.* Travaux du cercle linguistique de Prague, 7. (Appears in English translation as Principles of Phonology, translated by C. Baltaxe. Berkeley and Los Angeles: University of California Press, 1969).

VASSEUR, GASTON. 1963. *Dictionnaire des parlers picards du vimeu (Somme).* Collection de la Société de Linguistique picarde.

WARDHAUGH, RONALD. 1994. *An Introduction to Sociolinguistics.* 2nd ed. Oxford: Basil Blackwell.

YOUNG, RICHARD & ROBERT BAYLEY. 1996. VARBRUL analysis for second language acquisition research. In *Second Language Acquisition and Linguistic Variation,* Robert Bayley and Dennis Preston (eds.). John Benjamins Publishing Company, Amsterdam/ Philadelphia. 253-306.

ZWICKY, ARNOLD. 1977. *On Clitics.* Ms., IULC.

Bilingualism in Contemporary Finland: Whither Swedish?*

Sheila Embleton
York University

ABSTRACT

Bishop Henry of Uppsala led an expedition to Finland in the 12th century, beginning the strong connection of Finland to Sweden, which has since undergone various metamorphoses. Sweden ruled Finland until 1809, when Finland became a Grand Duchy of Russia. Finland's degree of autonomy varied over the years, until Finnish independence in 1917. An 1863 decree provided that Swedish would remain an official language, but that Finnish would gradually be introduced into all domains, becoming fully official by 1883. Currently Finland is officially bilingual, with complex rules as to what services must be provided in which languages in which communes. The official high status of Swedish masks some striking demographic, practical, and sociopolitical realities. The percentage of Swedish speakers has fallen rapidly over the years (now about 5%), and, coupled with the fact that Finnish Swedish has many dialects (all highly divergent from standard Swedish), this has led to practical consequences (to be detailed), which have in turn further diminished the role, status, and number of speakers. Recent developments in culture (internationalization, communications, etc.) and politics (decline in importance of the Nordic Council, sudden rise in importance of the European Community, rapid changes in eastern Europe, rise of other minorities, etc.) have combined to significantly boost the position and importance of English, and, as I will argue in this paper, concomitantly to weaken the position of Swedish dramatically, official bilingualism and rhetoric to the contrary.

King Erik IX Jedvardsson of Sweden (later St. Erik), accompanied by English-born Bishop Henry of Uppsala (later St. Henry, the patron saint of Finland), led an expedition to Finland in the 12th century (c. 1157). This expedition[1] had as its main aim the conversion of the 'heathen' Finns to Christianity. In the first half of the 13th century, another Englishman, Thomas, became Bishop of Turku (at the very southwestern corner of Finland), and worked vigorously to expand his diocese eastwards. Thus began the strong connection of Finland to Sweden, which has since undergone various metamorphoses, "evolving slowly from intermittent missionary endeavors to a loosely linked medieval society, to centralized bureaucratic monarchy, to the rise and eventual decline of an ambitious Swedish empire around the Baltic" (McRae 1997:12). This was also in many ways the beginning of Finnish connections to the outside world (i.e., Europe),

* I would like to thank Pentti Anttila, Raimo Anttila, Juhani Härmä, Sinikka Paukku, and two people who wish to remain anonymous for helpful discussion and material assistance during the writing of this paper.

[1] This is sometimes called the First Crusade to Finland, in contrast to the Second Crusade (or Häme Crusade), which took place shortly before the expedition to the Neva, "to protect the main route of expansion along the Gulf of Finland coast" (Klinge 1994:31). The Second Crusade was the one in which Alexander, Prince of Novgorod, drove back the combined Swedish and Finnish forces at the River Neva in 1240, earning himself the epithet 'Nevskiy'.

as those who managed to advance through the church hierarchy began to appear in the records of the universities of Paris, Prague, and various places in Germany[2]. Although Swedish political influence varied in intensity over the years, the influence of the Roman Catholic – and later Lutheran – church, from the west, remained a strong and constant presence against the influence of Novgorod/Russia and the Eastern Orthodox religion. In order to bolster the existing Swedish settlements and reinforce the Western influence, Swedish immigrants were encouraged into the region, particularly along the north coast of the Gulf of Finland (Helsinki/Uusimaa area, where it is visible even in the place-names, e.g. Helsinki < Hälsingland, the name of a region of Sweden). There was prolonged conflict over the boundary between the two spheres of influence, until the Treaty of Nöteborg/Pähkinäsaari/Schlüsselburg (now Petrokrepost, in Russia) was concluded between the Swedes and the Russians in 1323 (see Map 1a, from McRae 1997:14). Finnish/Western settlement gradually pushed further north and east of the line, and the boundary was formally altered a number of times to Sweden's benefit. Parts of Savo and Lapland, along with Estonia, became Swedish in 1595 (by the Treaty of Teusina/Täyssinä/Tyavzino; see Map 1a), and in 1617, by the Treaty of Stolbova, Sweden gained Kexholm/Käkisalmi (eastern Karelia) and Ingria (at the eastern end of the Gulf of Finland), and shortly thereafter acquired Livonia from Poland; (see Map 1a, and Map 2, from McRae 1997:18). It is important to realize that the spread of Western religion to Finland (as in other parts of the Baltic) "did not merely mean converting heathens but also preaching the Western doctrine and bringing the Western church organization" (Klinge 1994:22), particularly political organizational structure and taxes, to these areas. McRae (1997:20) also makes a particularly insightful comment, which I will quote at some length:

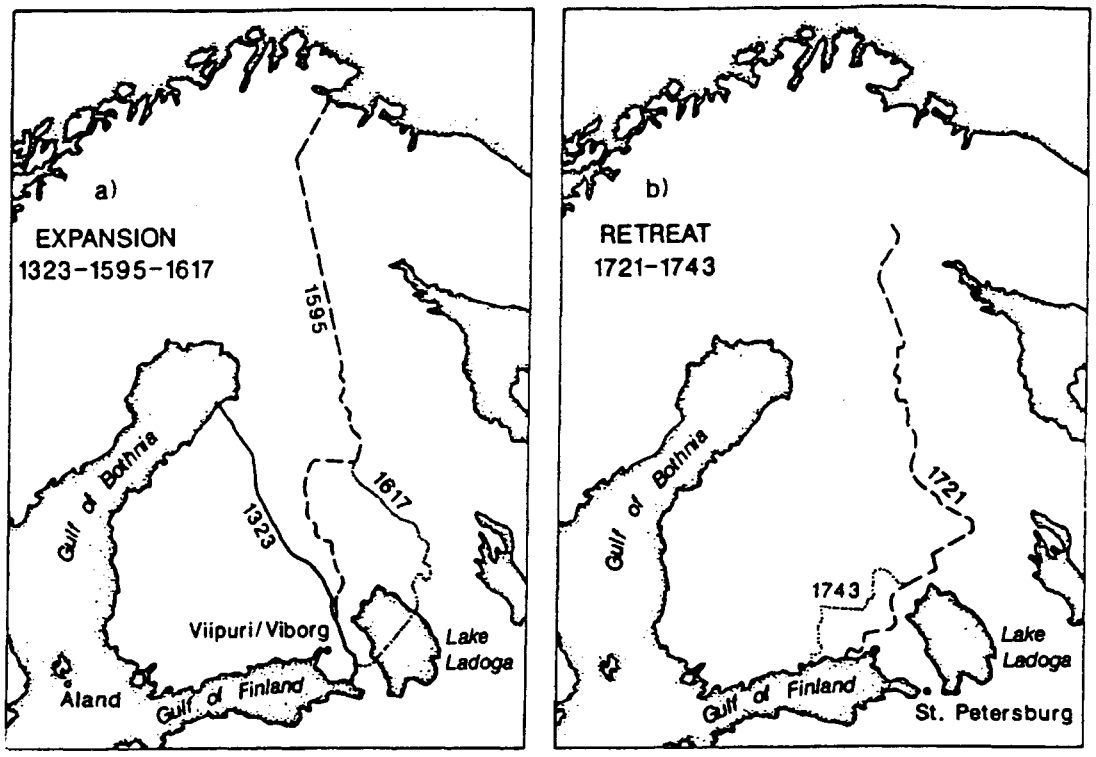
> For Finland the eighteenth century brought repeated military and diplomatic disasters. The latter part of the Great Northern War brought Russian invasion and an eight-year occupation of unusual severity for the entire country. Many members of the various elites – nobles, clergy, and civil servants – took refuge in Sweden proper, leaving the peasantry to cope as best they could by capitulation or retreat to the forests. Between 1692 and 1726, as one old guidebook records, Finland 'was visited by such great and numerous calamities as are seldom recorded in the history of nations' [...]. These misfortunes, which tend to be exaggerated in older historiography and in folk memory, included famines and military losses abroad as well as Russian occupation, but the memory of the latter helped in the long term to condition Finnish mass attitudes towards the Russians in a way that the elites, who had largely escaped this experience, would not always comprehend.

Throughout the period of Swedish rule (see also Map 1b, from McRae 1997:15), Swedish became the language of government and was the home language of the upper classes, but Finnish was never suppressed. Finnish was widely used in churches, by the upper classes to communicate with the peasantry, and there were even unofficial translations of legal documents into Finnish.

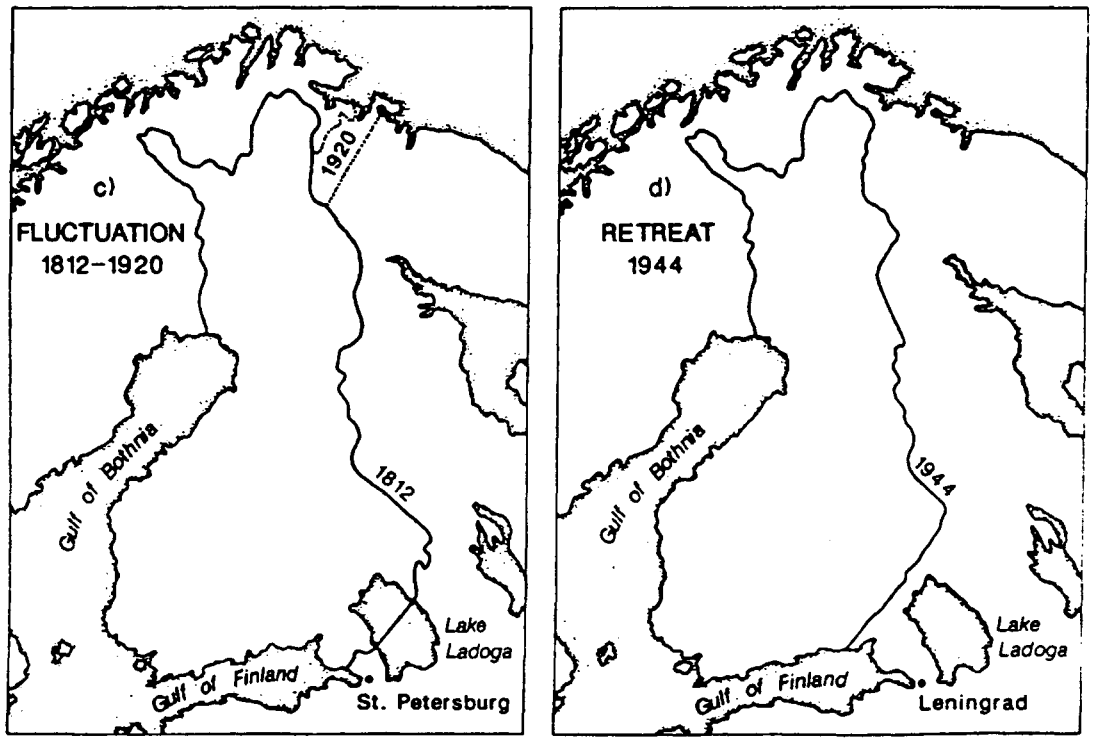
Sweden ruled Finland until 1809, when Finland became a Grand Duchy of Russia, at first under Tsar Alexander I. In trying "to wean Finland away from Sweden by making the connection

[2] Two Finns served as rectors of the University of Paris (Sorbonne) – Johannes Petri in 1366 and Olavus Magni from December 1435 to March 1436 (Heininen 1983:78; Palola 1998:13). Many other Finns served in various important positions at the University of Paris (see Heininen 1983 for details).

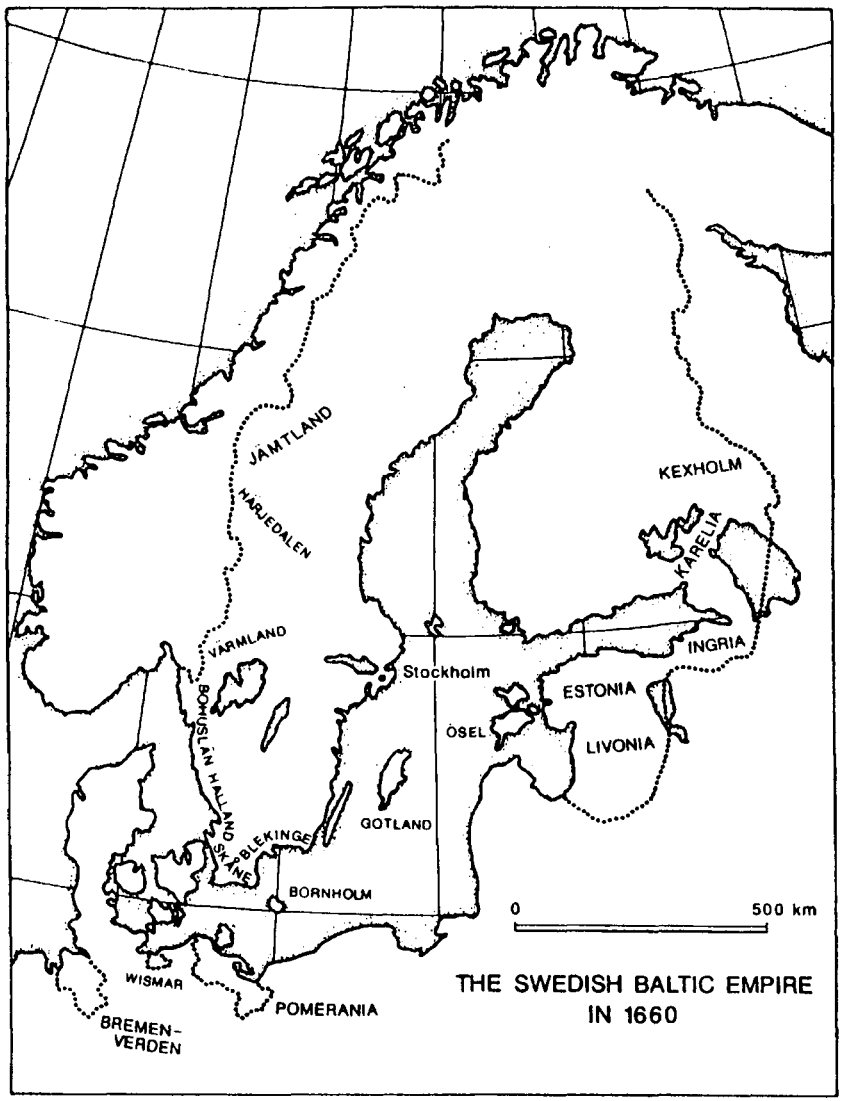
Map 1a-1b
The eastern frontier over seven centuries, 1323-1944



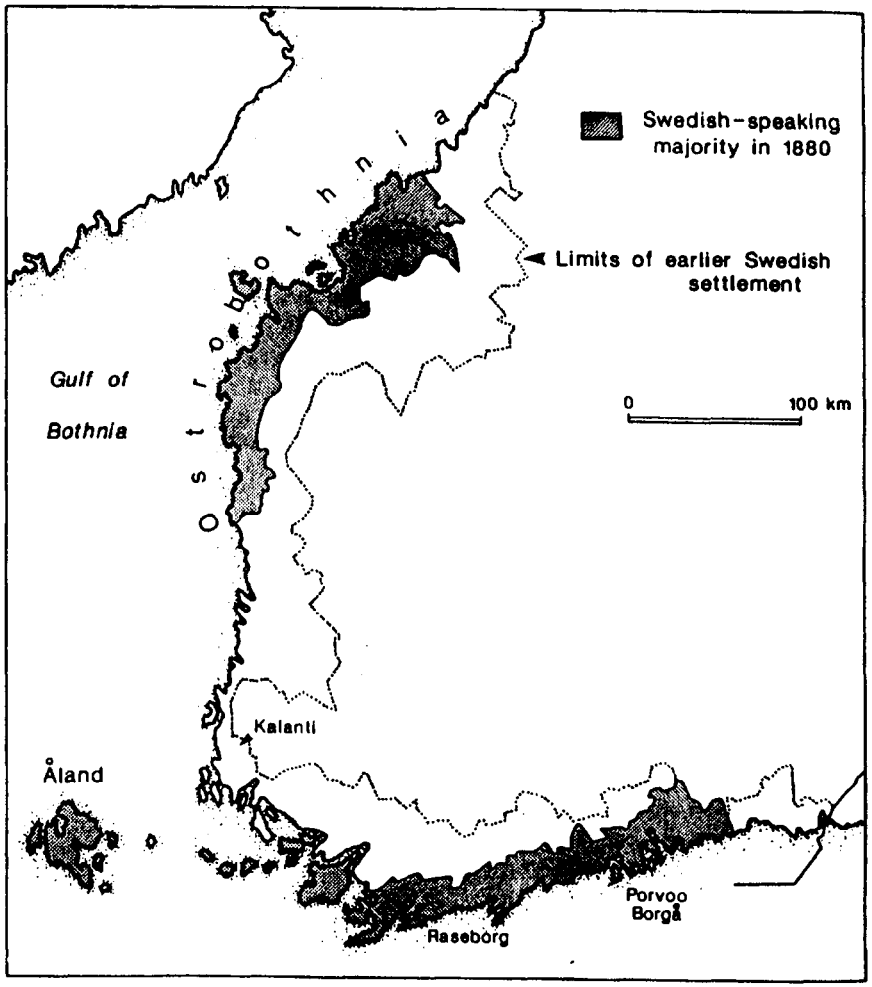
Map 1c-1d
The eastern frontier over seven centuries, 1323-1944—Continued



Map 2



Map 3
Swedish-speaking rural settlement, 1880 and earlier



with Russia more attractive and more profitable for all classes in Finland than the previous connection with Sweden" (McRae 1997:27), Finland was granted considerable autonomy – allowed to keep its own religion, legal system, Diet (parliament – Russia itself had no parliament at this time), education system, press, army units, postal system, currency, tariffs, trademarks, copyright, later railways, and most relevant for our purposes, language. During this period, Swedish and Russian were official languages, but Finnish remained in wide use (e.g., there were trilingual street-signs in Helsinki). The details are of course far more complex, but a nationalist (or national romantic) movement began to develop during the nineteenth century, along the lines of similar movements elsewhere in Europe and especially in Germany. In Finland, its three major figures were Elias Lönnrot (1802-1884; compiler of the *Kalevala*, Finland's national epic, first published in 1835, expanded and revised in 1849), J. L. Runeberg (1804-1877; Finland's first great national poet, and whose Swedish words are embodied in the Finnish national anthem), and J. V. Snellman (1806-1881; a social reformer who argued that the language of the masses, Finnish, should be the language of the nation)[3]. Along with other smaller victories, a breakthrough 1863 decree provided that Swedish would remain an official language, but that Finnish would gradually be introduced into all domains, becoming fully official by 1883. Again, McRae (1997:35) makes an insightful comment, slightly at odds with some prevailing Finnish views of this period:

> Russia was not unfavorable to Finnish cultural nationalism. Though the Tsarist authorities were suspicious of liberal tendencies, they were not against the Finnish language movement in principle. Anything tending to weaken Finland's connection with Sweden could be accounted in the Russian interest.

The pro-Finnish (Fennoman) movement was not without an antithetical pro-Swedish (Svecoman) movement, sometimes couched in terms of a pan-Scandinavianism. Finland's actual practical degree of autonomy from Russia varied over the years, but certainly took a downturn in the 1890s, particularly after the arrival of Governor Nikolai Ivanovich Bobrikov in 1898; Bobrikov set about a zealous Russification programme. One of the consequences for language was that Russian became the only official language, although both Swedish and Finnish were permitted in dealings between private citizens and government offices. The unpopular Bobrikov was assassinated in 1904. The national solidarity against Russification concretized more naturally behind Finnish than behind Swedish, as the language issue became increasingly bound up with other political issues. There are of course many twists and turns, many of them ironic and many of them still surprisingly relevant to the situation in the 1990s, both within Finland itself and concerning the relations between Finland and a new non-Soviet Russia. Finnish independence from Russia, in the wake of the Bolshevik Revolution, came on December 6, 1917. Unfortunately this was followed by a brief civil war in 1918, which left behind it a legacy of class-based conflict – although even here perceptions are different, with the "Reds" tending to see it as class-based and the "Whites" tending to see it as more of a war of liberation. This too is bound up in the language conflict, as there was a partial correlation between "Red", "lower-class", and "Finnish-speaking" as opposed to "White", "upper-class", and "Swedish-speaking", especially in some

[3] It is interesting to note that all three graduated from the University of Turku in the same year (1822). All three were native speakers of Swedish.

regions of the country. The final version of the constitution, signed in 1919, was relatively generous to the Swedish-speaking minority[4].

Article 14 of the 1919 Constitution (for much fuller and more detailed discussion, see McRae 1997:220ff.) has three separate paragraphs about language: that the national languages are Finnish and Swedish; that citizens have the right to deal with courts and administrative authorities equally in Finnish and in Swedish; and that cultural and economic needs of the Finnish-speaking and Swedish-speaking populations will be equally provided for[5]. The most important point to note is that it is absolutely equal status which is entrenched, not protection of Swedish-language minority rights. Article 22 provides that certain items must be promulgated in both official languages (e.g., laws, presidential decrees, etc.). Article 50 requires that in redrawing any provincial, district, commune, or municipal boundaries, as far as possible the boundaries should be drawn to produce monolingual units. This seems to be all that is left of some people's old aspirations in the 19th century for a type of self-government for each language group. Article 75 provides for linguistically homogeneous army units, but in a unique departure from the principle of absolute equality everywhere else, declares Finnish to be the language of command (presumably for purely practical reasons). Equally important as the 1919 Constitution itself are various other appended and related documents, e.g., those governing the language knowledge of civil servants, but particularly the Language Law of 1922, the most practically relevant parts of which will be summarized here. Each municipality or commune was classified as either unilingual (if the minority group is under 10% of the population) or bilingual (if the minority group is 10% or over); furthermore, Helsinki, Turku, and Vaasa would remain bilingual even if the minority would fall below 10%[6]. The classification was to be reviewed every 10 years, with the census, to reflect population shifts, and to avoid repeated changes in classification, there was to be a 2% "window" so that a commune would have to drop below 8% minority population to lose its bilingual status or rise above 12% minority status in order to become bilingual. Without going into details – and there are many – one of the most important effects is on the visibility of the minority language. In unilingual districts, messages, notices, road and street signs, etc. directed towards the general public are issued only in the one language, whereas in bilingual districts, all of these are supposed to be done in both languages (with the language of the numerically dominant

[4] For boundary changes 1812-1920, see Map 1c (from McRae 1997:15).

[5] Some of this is currently being rewritten, as a consequence of Finland's admission in 1995 to the European Union. The provisions have to be extended beyond just citizens to all residents – and since there will now be residents who do not speak either Finnish or Swedish, this should prove "interesting". Also, the article has recently been revised to recognize Saami rights, and to recognize the rights of the Romany and other groups to preserve and develop their language and culture (McRae 1997:220). There has already been one minor change, when the Nordic Languages Convention of 1981 was implemented in 1987. Citizens (not residents) of the Nordic countries (Finland, Sweden, Norway, Denmark, Iceland) have the right to use their mother tongue in virtually all dealings with authorities in any of the Nordic countries. This has had relatively little practical effect in Finland, since there have been so few Nordic immigrants (McRae 1997:230-231).

[6] Interestingly, all embassies and consulates abroad are defined as bilingual districts. It has been estimated that roughly 25% of the career diplomatic staff, and particularly at the highest levels, is mother tongue Swedish. Many diplomats also have acquired foreign spouses, which further complicates some situations. Off the record, I have been told that often only the language of the country of residence is used, e.g. in invitations, etc., particularly if it is a widely used language such as English, French, or German, precisely to avoid having to face certain linguistic questions head-on.

group first, and of the other group second). (Note "supposed to", as it is not always done – as but one example, during the summer of 1998 I personally saw many violations of this, for example with signage in train stations.) The internal working language (i.e., of the bureaucracy) within a district is obvious when it is a unilingual district, but is the language of the numerically stronger group in a bilingual district. Again, the details are complicated, because of the provision of certain types of services in the minority language no matter what, but this has a profound effect on for example employment opportunities for speakers of each language[7].

Since 1922, there have been a few changes. In 1935, during a wave of Finnish nationalism and linguistic tension[8], the guaranteed bilingual status of Helsinki, Turku, and Vaasa was removed, so they became subject to the same 10% rule as any other commune. But then the 1960 census contained somewhat of a shock – it showed that Swedish-speakers were only 7.6% of Turku's population, i.e. below the 8% threshold, and thus Turku, long the historic and statusful centre of Swedish life in Finland, was about to become unilingual Finnish. This occasioned some debate. In 1962, in a less tense linguistic climate – I would argue because Finnish speakers were now confident of their position – the Language Law was amended to state that any commune with a minority-language population of 5000 people (i.e., an absolute number, regardless of percentage) would become bilingual, and that the threshold to become bilingual was lowered to 10% from 12%. With the 2% "window", this meant that a commune could remain bilingual with only 8% minority population (or of course more than 5000 people from the minority group).

Notice how throughout all this, everything is still stated in scrupulously symmetric fashion. Swedish is not privileged, nor is Finnish. Everything is governed by percentages and absolute numbers, and applies equally, both ways around. The practical effects just may be different, given the overall demographic realities. For example, given this constraint of symmetry, the price to be paid for protecting Swedish in Turku could not be an outright and direct protection of Swedish in Turku, but had to be an overall lowering of the threshold everywhere for a commune to be considered bilingual – which indirectly threatened the unilingualism of some remaining Swedish communes. In 1975, a further formal legal amendment lowered the threshold for bilingualism from 5000 people to 3000 people, and to 8% from 10%; with the 2% "window", this means that a commune would have to slip below 6% before it would cease to be bilingual. The general tenor of the 1975 changes seemed to be to protect the rights of individual citizens, while lessening the overall blanket obligations of communes (and other organizations, e.g., hospitals, companies), likely in the face of the practical fact of decreasing Swedish language

[7] In 1991, Saami received essentially the same rights as Finnish and Swedish, which meant that four geographically large communes in Lapland (Inari, Utsjoki, Enontekiö, and part of Sodankylä) became bilingual (McRae 1997:231). This was immediately implemented for example in signage, which in many instances reflects two dialects of Saami, not just one. Strikingly, there are examples in Lapland where a sign is in Finnish, English, one or two dialects of Saami, and even German (for tourists), before Swedish will be added to the list.

[8] This was reflected in other ways too, such as many Finnish speakers changing Swedish-origin surnames into Finnish surnames, either by translation (e.g., *Stenberg* to *Kivimäki*, where Swedish *sten* = Finnish *kivi* = English 'stone' and Swedish *berg* = Finnish *mäki* = English 'hill') or by simply choosing a Finnish name, often a nature name (e.g., *Lindberg* to *Koskelo*, which is Finnish for 'merganser'). See Närhi 1987 for fuller detail.

competence in the population at large (both decreasing numbers of Swedish-speakers and decreasing numbers of Finns functionally competent in Swedish).

Currently Finland is officially bilingual, with complex rules as to what services must be provided in which languages in which communes, roughly along the lines described above[9]. But the official high status of Swedish masks some striking demographic, practical, and sociopolitical realities.

Demographics

The percentage of Swedish speakers has fallen rapidly over the years; it is now slightly over 5% (see Table 1, from McRae 1997:84).

Table 1

Finland: Total population by language, census years, 1865-1990 (percentages)

	Total population		Percentage speaking				
Year	(000s)	%	Finnish	Swedish	Sámi	Russian	Other and unknown
1865	1,843	100	85.7	13.9	0.03	0.2	0.1
1880	2,061	100	85.2	14.3	0.05	0.2	0.2
1890	2,380	100	86.1	13.6	0.05	0.2	0.1
1900	2,656	100	86.7	12.9	0.05	0.2	0.1
1910	2,943	100	88.0	11.6	0.06	0.3	0.1
1920	3,148	100	88.7	11.0	0.05	0.2	0.1
1930	3,463	100	89.4	10.1	0.06	0.2	0.2
1940	3,695	100	90.0	9.6	0.06	0.2	0.1
1950	4,030	100	91.1	8.6	0.06	0.1	0.1
1960	4,446	100	92.4	7.4	0.03	0.06	0.1
1970	4,598	100	93.2	6.6	0.05	0.04	0.1
1980	4,788	100	93.5	6.3	0.03	0.03	0.2
1990	4,998	100	93.5	5.9	0.03	n.a.	0.5

Sources: **1865-90**: *Befolkningsstatistik*, no. 29, 1899, 248-49; *Väestötilastoa*, no. 37, 1905, 68; **1900-80**: *Statistical Yearbook*, 1980, 42; 1992, 73.

In some ways, it is not so much the absolute number of Swedish speakers that has changed, as the number of Finnish speakers that has increased dramatically, thus affecting the relative percentages (see Table 2, from McRae 1997:86).

[9] I have left aside the question of Åland (Finnish Ahvenanmaa) entirely. Although Åland was assigned to Finland by a League of Nations decision in 1921, it is a 100% Swedish-speaking autonomous province, where the Finnish language has no status whatsoever, and in fact there are for example, restrictions on Finnish-speakers even owning property. For the history and details, see McRae (1997:61ff.) and Lundberg (1998)

Table 2

Finland: Swedish speakers and total population, numbers and percentages, selected years, 1610-1990

Year	Total population (000s)	Swedish speakers (000s)	Percentage change from previous decade	Percentage of total population
1610	n.a.	c. 70-73		c. 17-18
1749	n.a.	87		c. 16.3
1815	n.a.	c. 160		14.6
1840	1,446	c. 200		13.8
1865	1,843	256		13.9
1880	2,061	295		14.3
1890	2,380	323	+9.4	13.6
1900	2,713	350	+8.4	12.9
1910	2,921	339	−3.1	11.6
1920	3,105	341	+0.6	11.0
1930	3,381	343	+0.6	10.1
1940	3,696	354	+3.2	9.6
1950	4,030	348	−1.6	8.6
1960	4,446	331	−5.1	7.4
1970	4,287	303	−8.2	6.6
1980	4,788	300	−1.0	6.3
1990	4,998	297	−1.2	5.9

Sources: Fougstedt, 1951, 14; *Statistical Yearbook*, 1980, 42; Ekberg, 1982-85, 1: 341; *Finlandssvenskarna 1990*, 1992, 14.

There have also been two significant demographic losses of Swedish speakers. From the 1860s until 1914, about 330,000 people emigrated from Finland to North America, more than proportionately Swedish speakers, largely because Ostrobothnia, an area of heavy concentration of Swedish speakers, was over-represented in this emigration[10]. There was also a large post-war emigration to Sweden, for economic reasons, of about 240,000 people; for the years 1948 to 1950, these were approximately 47% Swedish speakers (McRae 1997:335). In the wake of the Winter (1939-40) and Continuation Wars (1941-44) in the early 1940s, Finland absorbed a huge number (approximately 400,000) of Karelian refugees[11], and these were more than proportionately Finnish-speaking and involved in agriculture. The Emergency Resettlement Law (June 1940) expropriated land to establish approximately 40,000 small farm-holdings; since "many of the larger farms from which land could be requisitioned were in Swedish areas" (McRae 1997:75), this raised linguistic concerns. After the second wave of Karelian refugees, 1945 legislation enunciated a general principle that resettlement should not cause any change in the classification of a commune as unilingual Swedish-speaking or bilingual; the informal guideline

[10] Total emigration up to 1930 is estimated at about 400,000, of which about 75,000 eventually returned to Finland. A rough estimate puts Swedish speakers at about 1/5 of the total in the earlier period up to 1914, and 1/3 of the total in the later period from 1924 to 1950 (McRae 1997:335, from various sources).

[11] For the boundary changes that resulted from these wars, see Map 1d (taken from McRae 1997:15).

used was no more than a 2% shift (McRae 1997:79)[12], [13]. It is also the case that Swedish speakers have a lower birth rate than Finnish speakers (McRae 1997:333).

Distribution

Swedish speakers are concentrated in four main areas; see Map 3, from McRae (1997:88). Since I am leaving Åland aside, as being a totally different situation, this leaves three areas: Helsinki and the Uusimaa/Nylånd area surrounding Helsinki, Turku and particularly its archipelago, and Ostrobothnia. The only significant change is urbanization, which causes both Swedish-speakers to move into towns (which are dominantly Finnish-speaking) and Finnish-speakers to move into towns (causing the absolute numbers of Finns in the towns to become higher). Thus the relative proportion of Finnish speakers to Swedish speakers becomes even higher, and Swedish-speakers are exposed even more than before to Finnish-speakers. Even in Helsinki, where there used to be a huge number of Swedish speakers, there were 19% Swedish-speakers in 1950, but only 7% Swedish speakers now; the current absolute number is about 37,000 (Tandefelt 1996:307-308). Helsinki has over 10% of the Swedish-speakers in Finland, the largest concentration, and so for this and other reasons is now the centre of Swedish-speaking life and culture in Finland. It is really the only place where one could even attempt to live a "full" life, including all levels and types of schooling and culture, in Swedish.

Practical Reality

All but the oldest most rural Swedish-speakers are essentially fully bilingual. Thus, e.g., almost all "service encounters" begin in Finnish (based on the probability of the native language of the interlocutor) and continue in Finnish (no reason to "notice" any accent or other deficiency in either interlocutor's use of Finnish, and hence no reason to switch from Finnish to Swedish). Mixed marriages usually result in children who are dominantly Finnish-speaking although reasonably competent in Swedish. This varies slightly in different social classes, in that higher social levels are more likely to maintain bilingualism to some degree (both at home and at work) and lower social levels are more likely to go over entirely to Finnish (both at home and at work) (Tandefelt 1996:309). The demographics are now such that there are so few Swedish-speaking schools that practical issues such as transportation become a concern, especially as both parents tend to be working. If the ratio of the number of Finnish-speakers to Swedish-speakers is roughly 95:5, i.e. 19:1, a rough estimate would give 19 Finnish schools for every 1 Swedish school. Thus the probability of a Finnish school being nearby, and a Swedish school being far away, is rather large. Furthermore,

> The bureaucracy in Finland has long had a public reputation for being less forthcoming in offering information to the public. By contrast with Canada or Belgium, for example, the linguistic rights of the public have not been well publicized, and it seems clear that many bilingual Swedish speakers are willing to deal with state or municipal employees in Finnish on grounds of efficiency and courtesy. ... In the

[12] Where possible, Swedish-speaking refugees were placed in Swedish-speaking or bilingual communes. Also, since Swedish-speaking farmers were to some extent protected from some of the burden of the expropriation due to the language law, they were assessed monetary payments instead, so that the state could purchase land (i.e., more land beyond the amount that could be expropriated) for Finnish-speaking refugees in Finnish-speaking communes.

[13] For great detail on the period up to the 1950s, see also Klövekorn (1960).

> absence of well-publicized norms, there seems to be ample room for daily practice to diverge from the theoretical model. (McRae 1997:229)

Translation costs have become astronomical. Already in the 1980s, there were 1500 unionized translators in Finland (McRae 1997:244). Swedish speakers seem content to receive translations of only the most important documents, or documents "on demand", rather than requiring routine translation of everything to which they might be legally entitled. Since almost all Swedish speakers in Finland – and certainly all educated speakers – can read Finnish fluently, there is no practical problem. Furthermore, since 1995, translation costs have risen yet again, after membership in the European Union[14]. A measure of the "contentedness" – or maybe the "resignation" – of Swedish speakers is the fact that the office of the Ombudsman [sic] receives only roughly 10 complaints per year involving language matters (out of about 1700). These tend to be quite minor, involving oversights or lack of resources, rather than any outright intentional breaches of the law (McRae 1997:263). It is of course impossible to separate cause and effect, but let us also look at second-language learning statistics in the schools. Within their 9-year school curriculum, all pupils must take 7 years of one second language (beginning in their third year of school), and 3 years of another second language (beginning in their seventh year of school). One of these must be the other national language (so Swedish or Finnish, as appropriate), and the other must be a modern international language. Two points are significant here. First, Finnish speakers have regularly argued that having to study two "foreign" languages (and it is interesting that this is how it is portrayed) is too burdensome. Second, the actual statistics (as of 1980) are that 87.4% of pupils in Swedish schools took Finnish as their first "foreign" language and 12.6% took English as their first "foreign" language. By contrast, in Finnish schools, 93.8% of pupils were taking English as their first "foreign" language, whereas only 4.9% were taking Swedish as their first "foreign" language, and 1.3% were taking other languages (mostly French and German) as their first "foreign" language (McRae 1997:298)[15]. If one looks at the statistics differently, namely which languages are studied as first, second, or third "foreign" language regardless of the declared mother tongue of the pupil, we find that English is studied by virtually 100% of the school population, German by 90%, French by 19%, Russian by 11%, Latin by 4%, and Spanish, Greek, and Saami combined by less than 1% (McRae 1997:298). Remember that 100% will be studying both Swedish and Finnish, but this "doesn't count", as it is a compulsory captive audience. Finally,

> In a setting of Nordic democratic values, the association of Swedish with a former dominant group and with a foreign power is a deterrent to linguistic activism in its defence. Swedish-speaking workers have been further cross-pressured by the earlier association of Swedish with the urban bourgeoisie and early capitalism. (McRae 1997:343)

---

[14] Not by any means the same thing, but various estimates give translation and interpretation costs within the European Community as roughly 40% of its annual budget.

[15] A third "foreign" language is also possible in the final two years, and of the few who choose this option, German predominates with 80.2%, then comes French with 12.4%, and Russian with 6.0% (McRae 1997:298). This situation has clearly changed (in favour of French, without necessarily diminishing German, after Finland joined the European Community in 1995).

Coupled with the fact that Finnish Swedish has many dialects (all highly divergent from standard Swedish, i.e. *rikssvenska*), this has led to practical consequences, which have in turn further diminished the role, status, and number of Swedish speakers. A generation ago, or even less, it would have been rare to find a Finn and a Swedish Swede (or other Scandinavian, for that matter) conversing in anything but Swedish (or rather some form of common Scandinavian, rather resemblant to Swedish) for example if they met at an academic conference, at some international political gathering, or via travel or tourism. But now such a conversation almost always takes place in English – to be sure the occasional words may be swapped in Swedish, particularly if there is some reason to reinforce some Nordic/Scandinavian solidarity, but the bulk of the serious conversation will be in English. This is increasingly the case even at primarily Nordic/Scandinavian gatherings, especially as other political factors have led such meetings to open out in their membership – particularly to no longer be restricted just to Norway, Denmark, Sweden and Finland, but also to include Iceland (with its far more divergent, albeit still "Scandinavian" language, Greenland, the Faroe Islands, Saami (from Norway/Sweden/Finland, and increasingly even Russia). Thus Swedish has become less important even in such settings, and English gradually becomes the lingua franca for practical purposes. As for the dialect differences to be found in Swedish within Finland itself, I will give just one example of problems of competing dialects within Finland Swedish. There was considerable controversy in the summer of 1998 about a Swedish-speaking news announcer on the national evening TV news in Swedish. She had an Ostrobothnian accent, and the Helsinki-based Swedish-speakers objected violently, with many saying they would rather not have the news in Swedish at all than with this accent[16].

There is an additional, but related and very practical, problem of which dialect of Swedish to teach to Finnish pupils in schools. Should it be the "local" Swedish dialect? Or Helsinki Swedish? Or *rikssvenska*? There is the question of local utility (and even pride) versus national utility, versus international utility; this practical but equally political problem should have a familiar echo for Canadians. There are problems all ways round, no matter which one is chosen. For example, using *finlandssvenska* in Scandinavia does surprisingly result in intelligibility problems as well as having a "country bumpkin" flavour; but using *rikssvenska* in Finland feels like affectation and even sounds humorous. This problem of choosing which form of Swedish of course then brings up the question of why Swedish at all, and not e.g. English.

There have been recent developments in culture and in politics which have combined to significantly boost the position and importance of English, and concomitantly to weaken the position of Swedish dramatically, official bilingualism and rhetoric to the contrary.

Developments affecting the cultural sphere include:
Internationalization and globalization of culture (for which English seems to be the world-wide vehicle);

[16] This type of seemingly irrational reaction occurs in other situations where languages are severely threatened, e.g. Surselvan/Romansh (Rhine Valley/Catholic) vs. Engadinish (Inn Valley/Protestant) in Rhaeto-Rumantsch, and even Acadian vs. Quebec dialects of French.

The dominance of US cultural "institutions" such as Hollywood and Disney (for which American English is obviously the vehicle), and the bulk of the music industry (thus teenagers know lyrics in English world-wide, even when they have no clue what those lyrics mean);
Regular, cable, and satellite TV. In Helsinki, in addition to the few Finnish channels (it is mostly a highly constrained state-operated system, YLE [Yleisradio]), which have a few minutes per day plus one whole evening per week in Finnish Swedish, there are currently three English-speaking stations available free and unscrambled – Eurosport, BBC World, and MTV (Music TV from England), as well as the French TV5, a Swedish station, a Russian station, and an Estonian station. If one subscribes to cable, then there are also several movie channels (which carry predominantly English-language material), CNN, Travel TV, Animal Planet, Cartoon, Discovery, Sky News, Sky Entertainment – all in English – as well as several other channels in European languages viewed as more "useful" than Swedish – Italian RAI Uno, Spanish TVE Internacional, and some material from Germany;
Other forms of more interactive communications (e.g., the rapid rise of the internet and web-pages, where most statistics show that Finns are more "connected" by internet and cell-phone than any other nation on earth – and the language of the internet is most decidedly English, and certainly not a language like Finnish that is full of diacritics);
Finnish travel and tourism abroad. Travel outside of Finland by Finns used to be more restricted to the Nordic area (where Swedish was useful), but now such travel is at least "Europe"-wide (Canary Islands, Greece, Turkey, etc.) and really worldwide (Thailand, North America, etc.). Thus English is the only language universally useful for travel in the places Finns go to these days.

Political developments include:
Internationalization and globalization in all aspects of life;
The decline in importance of the Nordic Council (not unrelated to the next point, about the rise in importance of European Union);
The fairly sudden rise in importance of the European Union. This brings us into a whole very complex – and rapidly emerging – set of questions. The de facto languages of wider communication within the European Union currently are English, French, and German, various political rhetorics to the contrary[17]. To the extent that other countries (Eastern European nations, Baltic nations, Russia, etc.) meet the criteria for membership and join, any further expansion is likely to lead to increased use of English, and decreased use of French and German, for complex historical, social, political, and linguistic reasons. In this McRae (1997:367) clearly agrees with me, when he says:

> the more the European Union supersedes the Nordic community as Finland's primary focus for external activity, the more the external advantage of Swedish is likely to cede to English, French, or German;

The rapid changes in eastern Europe, particularly in the Soviet Union;

[17] The working languages are officially English, French, and German (inter alia, McRae 1994:49-51), and this is clearly reflected in the job advertisements for European Union jobs, and (indirectly) in the demand for language training in the Finnish universities.

The rise of other minorities in Finland, both of domestic origin (i.e., Saami) and of neighbouring foreign origin (e.g., Estonian, Russian, Karelian, Ingrian; this is not unrelated to the previous point concerning the rapid changes in the former Soviet Union), as well as of more distant foreign origin (e.g., Vietnamese, Somali). If these people come knowing anything other than their own language, it will be English (certainly not Swedish or Finnish); they are only likely to learn Finnish (and maybe English, as they are often employed in service industries). They certainly have no time or energy to learn Swedish as well – and they do not feel any of those nationalistic or historical reasons to do so, only the basic practicality of having to learn a new language for daily life. Seen through their eyes, Swedish is not necessary, in any practical sense, for their daily lives.

The spinoff of some of the political developments mentioned above has been severe economic problems:
The 1990s saw economic problems in much of the developed world, but in Finland these were hugely exacerbated by the decline of the Soviet Union. There was the sudden and immediate loss of 25% of Finland's trade, with a concomitant rise in unemployment to over 25%. As often happens, there is a "salience of economic over linguistic issues in the public mind"; McRae (1997:159) already remarks on this with respect to the 1970s and 1980s, and it was only more so in the 1990s. Severe economic problems, where people's established and long-cherished life-styles are seriously threatened, in general tend to focus the mind on the practical and utilitarian aspects of the situation, rather than on history or idealism. English was seen to be practical and utilitarian in the modern world, whereas Swedish was seen as some sort of historical relic from the past, even if a link to the nation's history.

A new trend, certainly new within the past decade, has been the propensity in current advertising, to use much English, including incorrect English, to be "hip". An assortment of current examples is given in Appendix A. Another recent trend is for store-names to use English. For example, in Turku, there is a sports store called Pro Fish, and the shopping bags have "The Art of Fishing" written on them underneath the name, on both sides of the bag. An assortment of current examples is given in Appendix B. There is also a gradual increase in the use of English on signs in stores (and gradually less Swedish), and it is certainly not because there are suddenly more English-speaking tourists, because there are not and it occurs everywhere in the country, even where tourists of any kind are seldom seen. Often there is not even any sign in Finnish, just in English, which is a clear indication that this is not merely for some sort of practical value. There is also an increase in the use of English elsewhere in the media; for example, in August 1998, on the TV show "Hyvää huomenta Suomi" [Good Morning Finland], in the context of Daimler Benz choosing the Finnish GSM computer for its trucks, the announcer said "Siis suomalainen know-how" ('That's Finnish know-how'), and his interlocutor responded "Good". Perhaps his English response was prompted by the English loan-word *know-how*, but this sort of thing is ubiquitous nowadays. The English word *Yes* is also increasingly found in colloquial use, mostly in its "discourse" use, as opposed to its use as a direct affirmative.

Prognosis for Swedish

The genesis of this paper has been the change I have personally noticed over the past more than a dozen years of researching various aspects of language in Finland. I have also noticed an increasing tendency of various people – personal acquaintances, but also government bureaucrats – to outright lie about their language use. Reported and observed usage are often quite different. This has also alerted me to the fact that not only are changes taking place, but these changes are meaningful in some historical or other significant way – or else why lie about them? Something is clearly at stake. Nobody should be asked to predict the future – especially those of us who identify ourselves as historical linguists, and who have more than enough difficulty already predicting the past – but I would venture to say that if we were to fast-forward roughly 100 years, we would find that Finnish Swedish has all but disappeared. It would be interesting to see how the official rhetoric eventually deals with this – but of course none of us will be around to know the answer! In summary, then, this paper has had two main threads – the main one is the gradual demise of Swedish in Finland (and this is the rather weak pun of the title – whither and wither – as well as an echo back to a paper at a previous conference [Klinck 1996]), and the minor one is the concomitant rise of English, although obviously with a totally different function in the national psyche and soul.

REFERENCES

BEIJAR, KRISTINA, HENRIK EKBERG, SUSANNE ERIKSSON & MARIKA TANDEFELT. 1997. *Life in Two Languages – The Finnish experience.* Espoo, Finland: Schildts.

HEININEN, SIMO. 1983. Suomalaisten opinkäynti keskiajan Pariisissa [The Studying by Finns in Medieval Paris]. In Pentti Laasonen, Simo Heininen & Seppo J. Salminen (eds.), *Collegium Scientiae. (Suomen kirkkohistoriallisen seuran toimituksia,* 125.) Helsinki: Suomen kirkkohistoriallinen seura, 67-97.

KLINCK, ANNE. 1996. Whither Welsh? Strategies for survival. *Proceedings of the Twentieth Annual Meeting of the Atlantic Provinces Linguistic Association.* Fredericton, New Brunswick, 36-41.

KLINGE, MATTI. 1987. *A Brief History of Finland.* Helsinki: Otava. [Fifth edition, first edition is 1981]

KLINGE, MATTI. 1990. *Let Us Be Finns – Essays on history.* Helsinki: Otava.

KLINGE, MATTI. 1994. *The Baltic World.* Helsinki: Otava.

KLÖVEKORN, MARTIN. 1960. *Die sprachliche Struktur Finnlands 1880-1950. Veränderungen im sprachlichen Charakter der finnlandschwedischen Gebiete und deren bevölkerungs-, wirtschafts- und sozialgeographische Ursachen.* Helsinki: Centraltryckeriet.

LUNDBERG, ERIK. The Åland Islands – Coexistence with conflict? Talk given at University of Toronto, November 10, 1998.

McRAE, KENNETH D. 1994. The Changing International Environment and the Linguistic Stability of States and Regions: The European Community as a case study. *Journal of Behavioral and Social Sciences* (Japan) 3: 45-62.

McRAE, KENNETH D. 1997. *Conflict and Compromise in Multilingual Societies. Volume 3, Finland.* Waterloo, Ontario: Wilfrid Laurier University Press.

NÄRHI, EEVA MARIA. 1987. The Changing of Surnames in Finland during the Twentieth Century, Particularly in 1906 and 1935. *Studia Anthroponymica Scandinavica* 5: 109-121.

PALOLA, ARI-PEKKA. 1998. Yliopisto-opiskelua keskiajalla – Olavi Maununpoika Pariisissa [University Study in the Middle Ages – Olavus Magnusson in Paris]. *Tieteessä tapahtuu* 3/1998: 9-15.

SUNDMAN, MARKETTA. 1998. Review of Tandefelt 1996. *Virittäjä* 2/1998: 307-311.

TANDEFELT, MARIKA. 1996. På vinst och förlust. Om tvåspråkighet och språkförlust i Helsingforsregionen. (Forskningsrapporter från Svenska handelshögskolan, 35.) Helsinki: Svenska Handelshögskolan.

VIKØR, LARS S. 1993. *The Nordic Languages: Their status and interrelations.* Oslo: Novus Press & Nordic Language Secretariat.

APPENDIX A

Contemporary examples of English used in advertising in Finland

- *Meri Christmas / Viking Line.* Tram advertisement, November and December 1998, Helsinki. This is a pun, because Finnish *meri* means 'sea/ocean', so it is appropriate for a company advertising sea-transport and cruises.
- *Happy Tax Free 1999.* Tram advertisement for Viking Line, December 1998, Helsinki.
- Samuli Edelman's [a Finnish pop singer] newest record and CD is called *Greatest Hits.* TV advertisement, November 21, 1998, Helsinki.
- *Really cool, man!* First line in an advertisement from Merita bank for auto-leasing. *Helsingin Sanomat*, November 21, 1998.
- *The spirit of America on wheels.* First line in an advertisement for Chrysler cars and vans. *Helsingin Sanomat*, November 21, 1998.
- *Jeep. There's only one.* First line in an advertisement for Jeeps. *Helsingin Sanomat*, November 21, 1998.
- *Help wanted.* Graphic in an advertisement announcing five skilled positions to be filled by Sonera (the Finnish telephone company), *Helsingin Sanomat*, November 29, 1998.
- A mobile phone commercial for Sonera, where the screen has *Merry Xmas* (not Christmas), December 5, 1998.
- *Go before it's too late. The KILROY ticket. For young people under 26 and students under 33 only.* First line in a travel advertisement in *Helsingin Sanomat*, November 22, 1998. The same advertisement ends with *Check the KILROY site on www.kilroytravels.com.*
- *Eurovisit* as the headline on the price-list for all European destinations (all listed only with their Finnish names) and *Worldvisit* as the headline on the price-list for all the world-wide destinations (all listed only with their Finnish names*). Helsingin Sanomat*, December 6, 1998.
- *Explosive Skimbaaja.* Advertisement for the magazine Skimbaaja. *Helsingin Sanomat*, November/December 1998.

- *Look into our web pages. You may see your future.* This is the headline on an entire large job advertisement entirely in English, *Helsingin Sanomat*, December 1998.
- Lapin Kulta beer uses *The Golden Beer of Lapland* as its slogan.
- Nokia uses *Connecting People* as its slogan. This is worldwide, e.g., also found in the German magazine *Der Spiegel*, November 1998, along with a picture of two cell-phones and the caption *Perfect harmony.*
- A Nokia advertisement from December 1998 features a picture of a silver Nokia phone inside a gold-coloured chocolate box, with the *text Prepared with care from the finest ingredients.*
- An advertisement for Head skis contains the text *Enjoy the ride...* part-way through, *Helsingin Sanomat*, November 21, 1998.
- A take-out pizza counter in Helsinki (on Keskuskatu, in the Makkaratalo building) named *Take Away* advertises "*slice* kuin *slice*", i.e. that all slices of pizza cost the same, no matter what the toppings (December 1998).
- ASG, a transport logistics company, uses the slogan *C'est la vie – ASG myös tuo. Vie* is an eye-rhyme-like pun on *vie* (French 'life' and also Finnish 3rd person singular 'take away'); *myös tuo* means 'also brings' (Helsinki, December 1998).

APPENDIX B

Contemporary examples of English used in store-names, company-names, brand-names, etc. in Finland

- *After Dinner*, particular blend of Presidentti coffee [generally considered a fairly good or prestigious brand, somewhat more expensive, made by Paulig]. This advertisement is both in print and on television (e.g., on television in Helsinki, December 3, 1998, where "after dinner" is in fact also said).
- Travel agency Kohdematkat has as its slogan in its masthead *Leisure Travel, Helsingin Sanomat*, December 6, 1998 and elsewhere.
- Consulting company Vaisala has as its slogan in its masthead *Measure the Environment, Helsingin Sanomat*, December 13, 1998 and elsewhere.
- Construction company Koskisen has as its slogan in its masthead *stands for quality, Helsingin Sanomat*, December 1998 and elsewhere.
- Computer consulting company named *SiliconGraphics Computer Systems, Helsingin Sanomat*, December 1998 and elsewhere.
- Printing company named *WellPrint OY* [Ltd.], *with Green Valley Printing Department* also in its masthead. Job advertisement in *Helsingin Sanomat*, December 1998.
- A TV programme called *The Joulukalenteri* [Christmas calendar], December 5, 1998.
- A seafood restaurant at Helsinki-Vantaa airport called *The Mar Lodge*, seen December 1998 and January 1999. [*Mar* is the Spanish word for 'sea', which correlates with the theme of the restaurant, but the entire collocation seems quite peculiar as an English phrase. *Lodge* also seems inappropriate in the name of a small restaurant.]

The Seven Days and Their Names

MICHAEL FALK
Halifax

Introduction

The names of the week days in some languages go back to antiquity, while in others they have been introduced relatively recently. The present study was undertaken to elucidate the general conventions governing the formation of the day names and to examine the degree to which the world's languages follow these conventions. The language data for this article were collected from standard reference works, over 250 dictionaries in Halifax and Toronto libraries, as well as from the resources of the Internet.

The origin of the seven-day week

The moon cycle, on the average 29.5 days, gave rise to the lunar month, a measure of time that was universal in all cultures. However, the need to maintain regular market days and other recurring socio-economic or religious activities caused most early agricultural societies also to maintain shorter time cycles, which we could call "weeks". These early weeks consisted of three to twenty days in different cultures (Table 1) and may have originated in part because of a universal need for a day set aside for rest from heavy labour.

Table 1 . Some market-day cycles in the past.

Basque Culture (a)	3 days	
Bantu (b)	4 days	
Central Asia, Ireland (c)	5 days	
Babylonia (c)	7 days	
Rome (c)	8 days	(nundinae)
Inca (b)	8 days	
Greece (c)	10 days	(decades)
Egypt (b)	10 days	
Yoruba (b)	16 days	
Maya, Aztec (d)	20 days	

(a) Trask 1998. (b) Aveni 1989. (c) Calendars, Encyclopedia Britannica (1983).
(d) Stross 1983.

A few non-seven-day weeks are still in use today, for example an eight-day week in parts of Cameroon (Mopoho 1998). However, most of them were totally displaced by the current seven-day week, the original day names being largely forgotten. Nevertheless, we should not be surprised by an occasional survival of archaic day names in some of the world's languages today.

The seven-day week, now very nearly universal, appears to have originated in Babylonia between the eighth and sixth century BCE [O'Neil 1978]. The Babylonian month is known to have had special days, which included the 7th, 14th, 21st, and 28th. On these days travel was not to be undertaken and certain priestly functions, such as divination and healing, were not to be performed. The Babylonian month being strictly lunar, these days corresponded closely to the first quarter, the full-moon, the second quarter, and the disappearance of the moon (new moon). This must have been the beginning of the seven-day week, except that at the end of each four-week cycle there was either one or two extra days (Table 2).

Table 2. Babylonian lunar month.

1	2	3	4	5	6	7	◐
8	9	10	11	12	13	14	○
15	16	17	18	19	20	21	◑
22	23	24	25	26	27	28	●
29	(30)						

At some point, and we do not know exactly when, the extra days were left out, making the seven-day cycle continuous. This yielded the modern seven-day week, divorced from the lunar cycle. The seventh day of each cycle became known as the Sabbath (Hebrew *shabbat*, Aramaic *shabta*) that name probably derived from *shabattu*, the Akkadian name for the full moon [O'Neil 1978]. It is likely that the original meaning of *shabattu* became extended to include each of the four phases of the moon.

It should be stressed that the Hebrew root *sheva* 'seven' is unrelated to *shabbat*, and the Hebrew verb *shabbat* 'rested' is probably derived from *shabbat* 'the day of rest', rather than giving rise to it.

The numeric day names

When Jews adopted the seven-day week during their sojourn in Babylon in the sixth century BCE, they named the seven days numerically (Table 3). This is the system adopted in the Hebrew Bible. It represents the numeric convention of day naming in which the days are given the numbers one to seven (as in Genesis 1-2) but the seventh day may be given a special name, 'Sabbath' (as in Exodus 20).

Table 3. The day names in Genesis 1 and Exodus 20.

Hebrew	Aramaic	
Genesis 1		
Yom 'Exad	Yom'a Xad	'day one'
Yom Sheini	Yom Tinyan	'day second'
Yom Shlishi	Yom Tilitai	'day third'
Yom Rəviʃyi	Yom Rəviʃa'yi	'day fourth'
Yom Xamishi	Yom Xamisha'yi	'day fifth'
Yom Hashishi	Yom Shtita'yi	'day sixth'
Yom Hashəviʃyi	Yom'a Shəviyaʕ'ah	'day seventh'
Exodus 20		
Yom Hashabbat	Yom'a Shabta'	'the Sabbath day'

Transliterations used: y=י (yud) '=א (alef) x=ח (het) ʕ=ע (ayin) t=ת (tav) ə=schwa

Incidentally, there are two pecularities in the day numbering in Genesis 1-2, which do not appear to have been explained. (1) The numerals are all ordinal except for day one, for which a cardinal number is used. (2) The definite article is used for some of the days (sixth and seventh in Hebrew, first and seventh in Aramaic) and not for others. The name 'Sabbath' has passed on to very many of the world's languages, where it usually means 'Saturday', and occasionally 'Sunday', as indicated in Table 4.

Table 4. A few of the day names for Saturday, ultimately derived from Babylonian *shabattu*.

Europe:					
Catalan:	Dissabte	**Spanish:**	Sábado	**Italian:**	Sabato
French:	Samedi	**Sardo:**	Sappadu	**Rumanian:**	Sîmbătă
German:	Samstag	**Swabian:**	Samschdich	**Yiddish:**	Shabes
Russian:	Subbota	**Ukrainian:**	Subota	**Polish:**	Sobota
Czech:	Sobota	**Serbo-Croat:**	Subota	**Slovene:**	Sobota
Maced.:	Sabota	**Greek:**	Savvaton	**Chechen:**	Shot
Armenian:	Shapat	**Georgian:**	Shabati	**Ingush:**	Shoatta
Bizkaian:	Zapatu	**Hungarian:**	Szombat	**Maltese:**	Is Sibt
Asia:					
Arabic:	Al Sabt	**Farsi:**	Shambeh (c)	**Kyrghyz:**	Ishenbi (c)
Uzbek:	Shanba (c)	**Azeri:**	Senbe (c)	**Kazakh:**	Sembi (c)
Baluchi:	Shembe (c)	**Pashto:**	Shanba (c)	**Tajik:**	Shanbe (c)
Uyghur:	Shänbä (c)	**Turkmen:**	Shenbe (c)	**Bashkir:**	Shämbe (c)
Tatar:	Shimbä (c)	**Kurdish:**	Shemmé		
Africa:					
Harari:	Sabti	**Kabyle:**	Sebt	**Tuareg:**	Essebtin
Amharic:	Senbet (a)	**Guragi:**	Senbet (a)	**Oromo:**	Sambata (a)
Burji:	Sambata (b)	**Hausa:**	Subdu	**Mandinka:**	Sibitoo
Fula:	Aset	**Runyankore:**	Saabiiti (b)	**Bobangi:**	Sabala
Kanuri:	Sibda	**Teda:**	Essebdu	**Fulfulde:**	Assebdu
Begrimma:	Sibbedi	**Maba:**	Sab	**Malagasy:**	Asabotsy

(Table 4 continued)

Polynesia:

Marianese:	Sabalu	**Tagalog:**	Sabado	**Maranao:**	Sabtoo
Javanese:	Setu	**Malay:**	Sabtu	**Papua:**	Sabat (b)
Tongan:	Sapate (b)	**Majel:**	Jabot (b)		

Americas:

Paez:	Sapatu	**Tzotzil:**	Savaro	**Papiamentu:**	DjaSabra
Haiti:	Samedi	**Carrier:**	Sumdi	**Michif:**	Samjee

(a) Denotes both Saturday and Sunday. (b) Denotes Sunday. (c) Also means "week".

The planetary day names

An alternate naming convention, in which the seven days are named after the sun, the moon, and the five planets known in antiquity, also arose in Babylonia (Duncan 1988). The Babylonian names of the planets have been borrowed by many other cultures, including Greeks, Romans, Hindus, and Mongols, Babylonian gods being replaced by other, more or less equivalent, deities (Table 5).

Table 5. The names of the Divinities given to sun, moon, and the planets in antiquity.

	Babylonian Name (a)	**Latin Name**	**Greek Name**	**Sanskrit Name**	**Mongol Name** (b)
Sun	Shamash	Sol	Helios, Apollo	Surya	Miir
Moon	Sin	Luna, Diana	Selene, Artemis	Chandra,	Maag
Mars	Nergal	Mars	Ares	Angaraka, Mangala	(c)
Mercury	Nabu	Mercurius	Hermes	Budha	Tiir
Jupiter	Marduk	Iupiter	Zeus	Brihaspati,	Urmizt
Venus	Ishtar	Venus	Aphrodite	Shukra	Nagid
Saturn	Ninurta	Saturnus	Khronos	Shani	Kiwaan

(a) Duncan 1998. (b) Bazin 1989. (c) Not available.

The idea of a seven-day week with planetary day names had spread about the Mediterranean by the first century CE, probably driven by a popular intetest in astrology. The extent of this spread is attested by the bilingual graffiti, containing the Latin and Greek day names shown in Table 6, found in the excavation of Pompeii. These graffiti must have been scrawled during or before the year 79 CE, when Vesuvius erupted. Also, it is recorded that the Jews, who first appeared in Rome during the first century BCE, were referred to by the Romans as 'worshippers of Saturn'. This suggests that the day celebrated by the Jews as Sabbath coincided with the Roman *Dies Saturnis* (O'Neil 1978). By the time Christianity spread over the Roman Empire, the planetary day names were apparently well entrenched. Emperor Constantine legally incorporated the seven-day week into the Roman calendar in the year 321 CE, declaring *Dies Solis* an official day of rest and worship.

Table 6. Early planetary day names.

Planet	Pre-Christian Latin (79 CE)	Pre-Christian Greek (79 CE)	Sanskrit Name
Sun	Dies Solis	Heliu Hemera	Adityavaara or Ravivaara
Moon	Dies Lunae	Selenes Hemera	Somavaara
Mars	Dies Martis	Areos Hemera	Mangalavaara or Angarakavaara
Mercury	Dies Mercurii	Hermu Hemera	Budhavaara
Jupiter	Dies Iovis	Dios Hemera	Brihaspativaara or Curuvaara
Venus	Dies Veneris	Aphrodites Hemera	Shukravaara
Saturn	Dies Saturnis	Khronu Hemera	Shanivaara

Impact of Christianity

The attitude of the early Christian Church to the planetary names of the days is clear from the following three passages.

1. Ascribed to Saint Augustine. AD 354-430 (Dardel 1996)
   "*QUARTA FERIA* QUI MERCURII DICITUR A PAGANIS"

2. Ascribed to Caesarius, Bishop of Arles, Fifth Century (Holman1994)
   "NOS VERO, FRATRES, IPSA SORDISSIMA NOMINA DEDIGNEMUR...ET NUMQUAM DICAMUS DIEM MARTIS, DIEM MERCURII, DIEM IOVIS, SED *PRIMAM*, ET *SECUNDAM*, VEL *TERTIAM FERIAM*, SECUNDUM QUOD SCRIPTUM EST, NOMINEMUS..."

3. Ascribed to Pope Sylvester, AD 314-335 (O'Croinin 1981).
   "SANCTUS SILVESTER APOSTOLICUS SIC DOCUIT ET PRAEDICAVIT CHRISTIANIS, UT NON NOMINARENT DIES SEPTIMANAE IUXTA RITUM GENTILIUM, SED CHRISTIANA OBSERVATIONE SIC NOMINARENT, QUASI *FERIA PRIMA*, ID EST DOMINICUS, *SECUNDA FERIA, TERTIA FERIA, QUARTA FERIA, QUINTA FERIA, SEXTA FERIA, SEPTIMA FERIA*"

The early Church considered the planetary day names pagan and attempted to replace them by a numerical system (*prima feria, secunda feria, tertia feria*, ...) based on the Hebrew Bible. This Church terminology generally prevailed in Eastern Europe, except Finland. As Table 7 shows, the planetary names disappeared in Greek. They have been replaced by numeric names for Monday through Thursday, and religion-related names for Friday, Saturday, and Sunday.

Table 7. Impact of Christianity on Greek day names.

	Pre-Christan Greek	Modern Greek	
Sunday	Heliu	Kyriake	'Lord's day'
Monday	Selenes	Deftera	'2'
Tuesday	Areos	Triti	'3'
Wednesday	Hermu	Tetarti	'4'
Thursday	Dios	Pempti	'5'
Friday	Aphrodites	Paraskevi	'preparation'
Saturday	Khronu	Savvaton	from *Shabbat*

By contrast, the impact of Christianity in Western Europe was relatively minor. Romance languages retained five of the seven planetary names and adopted Church-related names only for Saturday and Sunday (Table 8). One must note one exception, however: the Church-related names were wholly adopted in Portuguese, replacing all seven planetary names. It is not clear why the Church was so uniquely successful in Portugal.

Table 8. Impact of Christianity on Latin day names.

Pre-Christian Latin	Church Usage	Medieval Latin	Modern Spanish	Modern Portuguese
Dies Solis	Dominica	Dominica #	Domingo #	Domingo #
Dies Lunae	Secunda Feria	Lunae (Lunis)	Lunes	Segunda-feira #
Dies Martis	Tertia Feria	Martis	Martes	Têrça-feira #
Dies Mercurii	Quarta Feria	Mercurii (Mercuris)	Miércoles	Quarta-feira #
Dies Iovis	Quinta Feria	Iovis	Jueves	Quinta-feira #
Dies Veneris	Sexta Feria	Veneris	Viernes	Sexta-feira #
Dies Saturnis	Sabbatum	Sabbata #	Sábado #	Sábado #

Names showing Church influence.

Later developments in Romance languages

Romance languages developed a three-fold split in the formation of the day names, shown by the three parts of Table 9.

Table 9. Monday to Friday in Romance languages.

Latin Type	Modern Language					
1. Dies Martis	**Catalan**	Diluns	Dimarts	Dimecres	Dijous	Divendres
	Occitan	Dilun	Dimers	Dimercre	Dijous	Divendre
2.Martis Dies	**French**	Lundi	Mardi	Mercredi	Jeudi	Vendredi
	Italian	Lunedi	Martedi	Mercoledi	Giovedi	Venerdi
3. Martis	**Spanish**	Lunes	Martes	Miércoles	Jueves	Viernes
	Romanian	Luni	Marţi	Miercuri	Joi	Vineri
	Sardinian	Lunis	Martis	Merculis	Zobia	Venere

In classical Latin the word *dies* could be attached to the front of the day name (*Dies Martis*) or to the back of the day name (*Martis Dies*) or it could be left out (*Martis*). All three types are preserved in Romance languages today. The unusual geographic distribution of these three types, puzzling at first, was recently explained by Dardel (1996).

Day names in Germanic languages

Germanic languages adopted the planetary day names in pre-Christan or early Christian times. *Dies Solis* and *Dies Lunae* were simply translated as 'sun-day' and 'moon-day', while for the names of the other five days, Germanic deities were substituted for the Roman Gods. In Old English all seven days bore planetary names, while in Old High German and Old Norse only six days did, the exception being Saturday, which was replaced at an early date by *Sambaztag* (from Greek *Sambaton*) and *Laugardagr* (bath-day), respectively (Table 10).

Table 10. Early Germanic day names.

Pre-Christian Latin	Old High German	Old English	Old Norse
Solis	Sunnūntag	Sunnandaeg	Sunnundagr
Lunae	Mānetag	Mónandaeg	Mánadagr
Martis	Ziestag	Tiwesdaeg	Tysdagr
Mercurii	Wodenstag	Wódnesdaeg	Óðensdagr
Jovis	Donerestag	Thunresdaeg	Thorsdagr
Veneris	Friatag	Frigedaeg	Friádagr
Saturnis	Sambaztag	Saternesdaeg	Laugardagr

Later, under Church influence, in the tenth century, German *Wodenstag* was replaced by *Mittawecha*, which became *Mittwoch*, while in Icelandic only the names of Sunday and Monday were retained, the other five day names being replaced by numbers or by religion-related terms

(Table 11). The introduction of *Mittwoch* and *Miðvikudagur* for Wednesday closely follows the popular late Latin *Media Hebdoma*, still found regionally as Tuscan *Mezzèdima*, Dolomite *Mesalèdema*, and Dalmatian *Misedma* (Holman 1994). Names modelled on *Media Hebdoma* and *Mittwoch* were coined in all Slavic languages, as well as in Finnish and Estonian. In Southern German dialects the name of Saturday remained *Samstag*, but in Northern dialects it was replaced by *Sonnabend* ('sun-eve'), which became standard. The day names in Norwegian and Danish are nearly the same as in Swedish (Table 11).

Table 11. Later developments in Germanic day names.

Dutch	German		Swedish	Icelandic	
Zontag	Sonntag		Sõndag	Sunnudagur	
Maandag	Montag		Måndag	Mánudagur	
Dinsdag	Dienstag		Tisdag	þriðjudagur #	'third-day'
Woensdag	Mittwoch #	'mid-week'	Onsdag	Miðvikudagur #	'mid-week-day'
Donderdag	Donnerstag		Torsdag	Fimmtudagur #	'fifth-day'
Vrijdag	Freitag		Fredag	Föstudagur #	'fast-day'
Zaterdag	Sonnabend # (or Samstag) #	'sun-eve'	Lõrdag #	Laugardagur #	'bath-day'

Non-planetary name.

Celtic languages

Subjected to early Romanization, Breton and Welsh borrowed all seven of the Roman planetary names. Apparently so did Cornish, although the data (Table 12) are incomplete. Several centuries later, the Scots, Irish, and Manx borrowed only three of the seven planetary names (Monday, Tuesday, and Saturday), adopting religion-related terms for the other four days.

Table 12. The day names in Celtic languages.

	Brythonic			Goidelic		
Latin	Breton	Welsh	Cornish	Irish Gaelic	Scots Gaelic	Manx
DiesSolis	Sul	DyddSul	Sül	AnDomhnach(b)	Di-Domnaich(b)	Jy-doone(f)
DiesLunae	Lun	DyddLlun	(a)	AnLuan	Di-Luain	Jyluain
DiesMartis	Meurz	DyddMawrth	Merth	AnMhàirt	Di-Màirt	Jemayrt
DiesMercurii	Marker	DyddMercher	Mergher	AnChéadaoin(c)	Di-Ciadaoin(c)	Jecrean(c)
Dieslovis	Diryaou	Dyddlau	(a)	AnDéardaoin(d)	Di-Ardaoin(d)	Jerdein(d)
DiesVeneris	Gwener	DyddGwener	Gwener	AnAoine(e)	Di-Haoine(e)	Jiheine(e)
DiesSaturnis	Sadorn	DyddSadwrn	Sadorn	AnSatharn	Di-Sathurn	Jysarn

(a) Not available. (b) From *Dominica*. (c) 'day of first fast'. (d) Etymology uncertain.
(e) 'day of fast'. (f) 'day of rest'.

Balto-Slavic languages

The Slavs were largely Christianized between 863 (Rotislav of Moravia) and 988 (Vladimir of Kiev). The Christianization resulted in a single set of names being spread over a wide area with a uniformity that is surprising, considering that the Slavs came partly under the control of the Western Church and partly under that of the Eastern Church. Table 13 shows the Slavonic day names today. This set of names was clearly coined initially under a strong Church influence. It contains special names for Saturday and Sunday, the remaining days being numbered. The name of Wednesday, *Środa* 'middle', clearly follows *Mittwoch, Miðvikudagur* and *Media-Hebdoma*. The persistence of this set of names in 11 languages for over 1000 years is remarkable, the only innovation having been the replacement of *Nyedyélya* by *Voskresyénye* (resurrection) in Russian. This replacement, which occurred around 1300, represents a substitution of one religion-related name by another.

Table 13. The day names in Slavonic languages.

Polish	Niedziela	Poniedziałek	Wtorek	Środa	Czwartek	Piątek	Sobota
Czech	Neděle	Pondělí	Úterý	Středa	Čtvrtek	Pátek	Sobota
Wend (Sorb)	Njedźela	Póńdźela	Wutora	Srjeda	Štwórtk	Pjatk	Sobota
Slovak	Nedel'a	Pondelok	Utorok	Streda	Štvrtok	Piatok	Sobota
Bulgarian	Nedelya	Ponedelnik	Vtornik	Sryada	Chetvyrtyk	Petyak	Sobota
Macedonian	Nedela	Ponedelnik	Vtornik	Sreda	Chetvrtok	Petok	Sabota
Serbo-Croat	Nedelja	Ponjedeljak	Utorak	Srijeda	četvrtak	Petak	Subota
Slovene	Nedelja	Ponedeljek	Torek	Sreda	četrtek	Petek	Sobota
Russian	Voskresyénye	Ponyedyélnik	Vtórnik	Sredá	četvérg	Pyátnitsa	Subbóta
Ukrainian	Nedilya	Ponedilok	Vivtorok	Sereda	Chetver	Pyatnitsa	Subota
Belarus	Niadziela ('no work')	Paniadziełak ('after-niedziela')	Awtorak ('second')	Sierada ('middle')	Chatvier ('fourth')	Piatnica ('fifth')	Subota (Latin *Sabbata)*

The Balts were Christianized after the Slavs, between 1259 and 1385. The Baltic day names (Table 14) are distinct from the Slavonic names, being entirely numeric, except for Sunday, which is called 'holy day'. The Balts have no equivalent of *Mittwoch*.

Table 14. The day names in Baltic languages

Lithuanian	**Latvian**	
Sekmādienis	Svētdiena	'holy day'
Pirmādienis	Pirmdiena	'first day'
Antrādienis	Otrdiena	'second day'
Trečiādienis	Trešdiena	'third day'
Ketvirtādienis	Ceturtdiena	'fourth day'
Penktādienis	Piektadiena	'fifth day'
Seštādienis	Sestdiena	'sixth day'

Estonian and Finnish

Though these two languages belong to the Ugro-Finn family and are very closely related, they have very different systems of day names. Finnish appears to have simply borrowed the Scandinavian planetary names, while Estonian seems to have borrowed only the names of Friday and Saturday the other days being named in a numeric system that recalls the Slavic model (Table 15).

Table 15. The day names in Estonian and Finnish.

Old Norse	Finnish	Estonian	
Sunnundagr #	Sunnuntai #	Pühapäev	*püha* = 'holy', *päev* = 'day'
Mánadagr #	Maanantai #	Esmaspäev	'first-day'
Tysdagr #	Tiistai #	Teisipäev	'second-day'
Óoendagr #	Keskiviikko	Kesknädal	'mid-week'
Thorsdagr #	Torstai #	Neljapäev	'fourth-day'
Friadagr #	Perjântai #	Reede #	from *Friadagr*
Laugardagr	Lauantai	Laupäev	'bath-day'

Name of planetary origin

Hungarian, Turkish, Basque, and Albanian

All four of these languages appear to preserve some archaic forms of day names.

In Hungarian (Table 16) the names of Wednesday though Saturday are borrowed from the surrounding Slavonic languages, but the names Vasárnap (Sunday), Hétfö (Monday), and Kedd (Tuesday) are non-Slavonic. Vasárnap means 'market day' and may be a loan-borrowing from Turkish, but Hétfö (week-head) and Kedd (etym. dub.) may be of pre-Christian origin.

Table 16. The day names in Hungarian and Turkish.

Hungarian		Turkish	
Vasárnap	'market-day'	Pazar	'market'
Hétfö	'week-head)'	Pazartesi	'after-market'
Kedd	#	Sali	#
Szerda	'middle' (Slav)	Çarşamba	'four-week' (Persian)
Csütörtök	'4' (Slav)	Perşembe	'five-week' (Persian)
Péntek	'5' (Slav)	Cuma	(Islamic)
Szombat	from *Sambaton*	Cumartesi	('after-*Cuma*')

Etymology uncertain.

Turkish day names (Table 16) include a Persian borrowing for Sunday (*Pazar* or market), numeric days for Wednesday and Thursday, also borrowed from Persian, and the Islamic *Cuma* for Friday. The names for Saturday and Monday, *Cumartesi* (after-*Cuma*) and *Pazartesi* (after-market), are apparent innovations, though the latter appears to follow the pattern of Slav

Poniedziałek (Table 13). The name for Tuesday, *Sali*, is of uncertain derivation (possibly from Arabic *Tsoulatsa?*), and, like the corresponding day name in Hungarian, may be archaic.

Basque day names (Table 17) are interesting. The name for Sunday, *Igande* is Church-related, but the other names could be archaic. The names for Monday, Tuesday, and Wednesday (*Astelehen, Astearte, Asteazken*) appear to imply a three-day week (Trask 1998).

Albanian day names (Table 17) are planetary for Saturday through Wednesday, but the names of Thursday and Friday, *Enjte* and *Prémte* could be archaic.

Table 17. The day names in Basque and Albanian.

Basque		**Albanian**	
Igande	'resurrection'	Diel	'sun'
Astelehen	'week-first'	Hënë	'moon'
Astearte	'week-middle'	Martë	'Mars'
Asteazken	'week-last'	Mërkurë	'Mercury'
Ortzegun	'sky-day'	Enjte	#
Ortzirale	'sky-?'	Prémte	#
Larunbat	#	Shtunë	'Saturn'

Etymology uncertain.

Day names in the languages of the Caucasus

These languages have no planetary names (Table 18). Instead, they have numeric names for Monday through Thursday, and Church-related names for Friday, Saturday, and Sunday, borrowed from Medieval Greek. For Monday, Chechen and Ingush have apparently borrowed Georgian *Orshabati*, or 'day two', but then they call Tuesday *Shinara* which also means 'two' in their own language. There are many examples of this type of confusion involving two counting systems.

Table 18. The day names in the languages of the Caucasus.

Armenian	**Georgian**	**Chechen**	**Ingush**	
Giragi	K'wira	K'irande	K'irandi	'*Kyriake' (Greek)*
Yergushapti '2'	Orshabati '2'	Orshot '2'	Oarshuot '2'	
Yerekshapti '3'	Samshabati '3'	Shinara '2'	Shinara '2'	
Chorekshapti '4'	Otxshabati '4'	Qaara '3'	Qeara '3'	
Hinkshapti '5'	Xutshabati '5'	'Eara '4'	Jiera '4'	
Urpat #	P'arask'evi	P'eraska	Ruzba #	'*Paraskevi*' (Greek)
Shapat	Shabati	Shot	Shoatta	'*Sabbaton*' (Greek)

Etymology uncertain.

The impact of Islam

Under Islam, the all-important day of the week became the sixth day, Friday, which has been named *Juma'a*, meaning 'assembly'. In addition, Islam has borrowed the name of Sabbath for the seventh day of the week, *Al Sabt* in Arabic. For the other days the Arabs adopted the Hebrew numeric approach (Table 19).

Table 19. The day names in modern Arabic.

Day	Name	
Sunday:	Yaum Al-Axad	'day-one'
Monday:	Yaum Al-Itsnain	'day-two'
Tuesday:	Yaum At-Tsulatsa	'day-three'
Wednesday	Yaum Al-Arba'a	'day-four'
Thursday:	Yaum Al-Xamiis	'day-five'
Friday:	Yaum Al-Juma'a	'day of assembly'
Saturday	Yaum Al-Sabt	'day of Sabbath'

In many languages in the Islamic world, the day names have been borrowed from Arabic (Table 20).

Table 20. Some numeric day names borrowed from Arabic.

Language	Sunday(1)	Monday(2)	Tuesday(3)	Wednesday(4)	Thursday(5)
Arabic	Al-Axad	Al-Itsnain	Al-Tsoulatsa	Al-Arbaa	Al-Xamiis
Ladino (Spain)	Alxad	(a)	(a)	(a)	(a)
Maltese	Il-Ħadd	It-Tnejn	It-Tlieta	L-Erbgħa	Il-Ħamis
Harari (Ethiopia)	Alxad	Isniin	Säläsa	Arba'a	Xamiiš
Somali	Axad	Isniin	Talaado	Arbaco	Xamiis
Tuareg (Sahara)	Elxedden	Lîtniten	Ettenâtetîn	Inardâten	Elremîsen
Kabyle (Algeria)	Elxad	Tnain	Tlata	Elarbâa	Xmis
Amharic (Ethiopia)	Ixud	Senio	Makseniu (b)	Rebuu	Xamus
Hausa (Nigeria)	Lahadi	Lìtìnîn	Tàlata	Laraba	Alhamis
Bahasa Malasia	Ahad	Isnin	Selasa	Rabu	Kamis
Maranao (Phil.)	Akad	Isnin	Salasa	Arbaqa	Hamis
Indonesian	Ngahad	Senèn	Selôsô	Rebo	Kemés
Malagasy (Madag.)	Alahady	Alatsinainy	Atalata	Alarobia	Alakamisy
Mandinka (Gambia)	Alahadoo	Tenan	Talatoo	Araboo	Araamisoo

(a) Non-Arabic name. (b) 'after-*Senio*'.

However, not all Islamic day names are borrowed from Arabic. In modern Persian (Farsi) the only day name borrowed from Arabic is *Juma'a* (Friday). The other days are numbered in a system analogous to that in Armenian and Georgian (Table 18), using a numeral plus *shambeh*, a counter meaning 'week', borrowed from Greek *sambaton*. Table 21 shows some of the Persian

day names adopted by Indo-European languages closely related to Persian, like Kurdish, Baluchi, and Tajik, or by entirely unrelated Turkik languages like Uzbek, Kyrghyz, Uighur, Kazakh, Turkmen, Bashkir, and Tatar.

Table 21. Some numeric day names borrowed from Persian.

	Sunday (1)	Monday (2)	Tuesday (3)	Wednesday (4)	Thursday (5)
Indo-European Languages:					
Farsi	Yekshambeh	Doshambeh	Seshambeh	Chaharshambeh	Panjshambeh
Kurdish	Yekshemmé	Dushemmé	Seshemmé	Chwarshemmé	Penjshemmé
Baluchi	Yekshembe	Dwshembe	Seyshembe	Charshembe	Penchshembe
Tajik	Yakshanbe	Dushanbe	Seshanbe	Chorshanbe	Panjshanbe
Turkik Languages:					
Uzbek	Yakshanba	Dushanba	Seshanba	Chorshanba	Paishanba
Kyrghyz	Jekshenbi	Düyshünbü	Sheishenbi	Charshenbi	Beyshenbi
Uyghur	Yäkshänbä	Dushänbä	Sayshänbä	Charshänbä	Päyshänbä
Kazakh	Jekshembi	Düysembi	Seysembi	Särsembi	Beysembi
Turkmen	Ekshenbe	Düshenbe	Sishenbe	Charshenbe	Penshenbe
Bashkir	Yäkshämbe	Düshämbe	Shishämbe	Shärshambe	Kesadna #
Tatar	Yäkshämbe	Dushämbe	Sishämbe	Chärshämbe	Pänjshämbe

Non-Persian name.

Day names in the languages of India

The planetary day names spread to India in pre-Christian times. Variants of the Sanskrit names (*Adityavaara, Somavaara,...*), are used today in all the Indo-European languages of India, in many of the unrelated Dravidian languages, like Telugu and Tamil, and even in some Mon-Khmer languages of Indochina, including Cambodian, Lao, and Thai (Table 22).

Table 22. Planetary day names borrowed from Sanskrit.

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Sanskrit	Aditya or Ravi (Sun)	Soma (Moon)	Mangala Or Angaraka (Mars)	Budha (Mercury)	Brihaspati or Guru (Jupiter)	Shukra (Venus)	Shani (Saturn)
Hindi	Ravivaar	Somvaar	Mangalvaar	Budhvaar	Brihaspativaar	Shukravaar	Shanivaar
Marathi	Rawiwar	Somwar	Mangalwar	Budhwar	Gurwar	Shukrawar	Shaniwar
Bengali	Robibar	Shombar	Mongalbar	Budhbar	Brihaspatibar	Shukrabar	Shonibar
Assamese	Rabibar	Hombar	Mangalbar	Budhbar	Brihaspatibar	Hukurbar	Hanibar
Panjabi	Aitwaar	Somwaar	Mangalwaar	Buddhwaar	Wiirwaar	Shukkarwaar	Saniccar
Nepali	Raviwar	Som	Mangal	Budhabar	Brihaspati	Shukrawar	Shaniwar
Urdu	Itwaar	Piir #	Mangal	Budh	Jumaraat #	Juma #	Sanichar
Singhalese	Iridha	Sandudha	Angeharuvadha	Budhdha	Brihaspetindha	Sikuradha	Senesuradha
Telugu	Aadivaaram	Somavaaram	Mangalvaaram	Budhavaaram	Guruvaaram	Shukruvaaram	Sanivaaram
Cambodian	Tngay-Qaattit	Tngay-Chun	Tngay-Ong'kea	Tngay-Puut	Tngay-Prohoa	Tngay-Sok	Tngay-Saw
Lao	Wan-Aathit	Wan-Jan	Wan-Angkhan	Wan-Phut	Wan-Phahat	Wan-Suk	Wan-Sao
Thai	Wun-Ahtit	Wun-Jun	Wun-Umgkahn	Wun-Poot	Wun-Pareuhut	Wun-Sook	Wun-Sao

Non-planetary name.

Day names in Bantu languages

The Bantu languages generally borrowed the name of Sunday from English and developed a numeric system for the other six days, starting the count with Monday (Table 23). Swahili is an interesting exception. Under the Islamic influence, it named Friday *Ijumaa* and numbered the other days of the week, starting the count with Saturday, so that its numbering is at odds with that of the other Bantu languages. For Wednesday, Swahili has *Jumatano*, which contains the numeral *tano* (five). For Thursday, Swahili borrowed the Arabic name *Alhamisi*, so that it has two consecutive days bearing the number five, another confusion of day-counting systems.

Table 23. Numeric day names in some Bantu languages.

Shona (Zimbabwe)	**Zulu** (S. Africa)	**Bemba** (Zambia)	**Tonga** (N.Zimbabwe)	**Luganda** (Uganda)	**Swahili** (E. Africa)
Svondo (a)	iSonto (a)	Mulungu (b)	Nsondo (a)	Ssande (a)	Jumapili '2'
Muvhuro	uMsombuluko	Cimo '1'	Musumbuluko	Bbalaza (b)	Jumatatu '3'
Chipiri '2'	oLwesibili '2'	Cibili '2'	Bwabili '2'	Lwakuburi '2'	Jumanne '4'
Chitatu '3'	oLwesithatu '3'	Citatu '3'	Bwatatu '3'	Lwakusatu '3'	Jumatano '5'
China '4'	oLwesine '4'	Cine '4'	Bwane '4'	Lwakuna '4'	Alhamisi '5' (c)
Chisanu '5'	oLwesihlanu '5'	Cisano '5'	Bwasanu '5'	Lwakutaano '5'	Ijumaa (c)
Mugovera (b)	iMigqibelo (b)	Cibelushi (b)	Mujibelo (b)	Lwamukaaga '6'	Jumamosi '1'

(a) From 'Sunday'. (b) Etymology unknown. (c) From Arabic.

Tibetan and Chinese day names

Tibetan has preserved a complete set of planetary day names, while modern Chinese has replaced the planetary names with a numeric system of day naming (Table 24). However, Chinese preserves a planetary name for Sunday, 'sun-day' in Cantonese and 'sky-day' in Mandarin.

Table 24. The day names in Tibetan and Chinese

	Tibetan (a)		**Mandarin**	
Sunday	Nyi-ma	'sun-day'	Xing"qi"tian`	'week-sky'
Monday	Zla-wa	'crescent, moon'	Xing"qi"yi"	'week-1'
Tuesday	Mig-mar	'red-eye, Mars'	Xing"qi"er`	'week-2'
Wednesday	Lhag-pa	'Mercury'	Xing"qi"san"	'week-3'
Thursday	P'ur-bu	'Jupiter'	Xing"qi"si`	'week-4'
Friday	Pa-sangs	'Venus'	Xing"qi"wu^{v}	'week-5'
Saturday	Spen-pa	'Saturn'	Xing"qi"liu`	'week-6'

(a) Krueger 1962.

Japanese and Quechua day names

These two unrelated languages, half a world apart, independently adopted planetary day names (Table 25). The first two days are the same: sun-day and moon-day. The next five days continue differently, reflecting distinctive cultural traditions. In Japanese these days are named after the five elements that make up the physical world (fire—water—wood—metal—earth), while in Quechua they are named after other sky-related terms (wizard—star—Venus—lightning—rainbow). Some of these names may be archaic.

Table 25. Planetary day names in Japanese and Quechua.

Day	**Japanese**		**Quechua**	
Sunday	Nichiyoobi	'sun-day'	Intichay	'sun-day'
Monday	Getsuyoobi	'moon-day'	Killachay	'moon-day'
Tuesday	Kayoobi	'fire-day'	Atipachay	'wizard-day'
Wednesday	Suiyoobi	'water-day'	Qoyllurchay	'star-day'
Thursday	Mokuyoobi	'wood-day'	Ch'askachay	'Venus-day'
Friday	Kinyoobi	'gold-day' #	Illapachay	'lightning-day'
Saturday	Doyoobi	'earth-day'	K'uyichichay	'rainbow-day'

or 'metal-day'.

Day names in Amerindian and Inuit languages

Most of these languages adopted a numerical system of day names (Table 26).

Table 26. Numerical day names in some Amerindian languages.

	Guarani (Paraguay)	**Mohawk** (S.W. Quebec)	**Ojibwe** (Anisshinaabe, S.E. Ontario)	**Potawotomi** (Oklahoma)	**Inuktitut** (Arctic Bay, Baffin Island)
Sunday	Arateĩ (1)	Tsatahkhaton (7)	Anami'e (a)	Numagishguk (d)	Naatiunguya (f)
Monday	Arakõi (2)	Tiotierenhton (1)	Ishkwaa-anami'e (b)	Ngotgonguk (1)	Noggaja (c)
Tuesday	Araapy (3)	Tekenihaton (2)	Niizho (2)	(e)	Aippaat (2)
Wednesday	Ararundy (4)	Ahsenhaton (3)	Aabitose (c)	Nsognakuk (3)	Pingajuat (3)
Thursday	Arapo (5)	Kaierihaton (4)	Niiyo (4)	Yawkokak (4)	Qitingquut (4)
Friday	Arapoteĩ (6)	Wiskhaton (5)	Naano (5)	(e)	Ullutuinnaq (g)
Saturday	Arapokõĩ (7)	Iahia:khaton (6)	Giziibiigisaginige (c)	(e)	Sivataarvik (h)

(a) 'week'. (b) 'after-*Anami'e*'. (c) Etymology uncertain. (d) *nume* ='church'. (e) Not available. (f) 'the day to do nothing. (g) 'just any old day'. (h) 'the day to receive biscuits'.

The count starts with Sunday in some cases and with Monday in others. In a few of the languages, for example in Guarani and Mohawk, all seven of the days bear numeral names, but more commonly the days bear a mix of numerical and non-numerical names. Some of the non-numeric names are religion-based, but some are very descriptive, as in Inuktikut (Table 26) or in

Blackfoot, in which the names for Saturday and Wednesday translate as 'small shopping day' and 'the day when we have already drawn our rations'.

Day names borrowed from other languages

In many languages the numeric or planetary day names have been borrowed without any apparent understanding of their original meaning. The names that have been borrowed most widely are Arabic, Persian, English (Table 27), Spanish (Table 28) and French (Table 29). Sometimes all seven day names have been borrowed, as in Majel, Michif, or Tzotzil. In other cases only some of the day names have been borrowed, native names having been developed for the remaining days, as in Tongan, Carrier, or Maori.

Table 27. Some day names borrowed from English.

Papua-Pidgin Torres-Strait	**Papua-Pidgin** Port-Moresby	**Tongan**	**Majel** (Marshall Islands)	**Maori** (NZ)	
Sande	Sande	Sapate	Jabot	Ratapu #	'holy-day'
Mande	Mande	Monite	Manre	Mane	
Tyuzde	Tunde	Tusite	Juje	Turei	
Wenezde	Trinde	Pulelulu #	Wonje	Wenerei	
Tazde	Fonde	Tuapulelulu #	Taije	Taite	
Praide	Fraide	Falaite	Balaire	Paraire	
Satade	Sarere	Tokonaki #	Jarere	Rahoroi #	'clean-day'

Name not borrowed from English.

Table 28. Some day names borrowed from Spanish.

Tzotzil (Mexico)	**Papago-Pima** (Arizona)	**Papiamentu** (Curaçao)	**Chamorro** (Marianas)	**Tagalog** (Philippines)
Rominko	Domig	Djadumingu	Damenggo	Linggo (a)
Luneš	Luhnas	Djaluna	Lunes	Lunes
Marteš	Mahltis	Djamars	Mattes	Martes
Melkukeš	Mialklos	Djarason (a)	Metkoles	Miyerkules
Hweveš	Huiwis	Djaweps	Huebes	Huwebes
Byerneš	Wialos	Djabierne	Betnes	Biyernes
Savaro	Shawai	Djasabra	Sabalu	Sabado

(a) Name not borrowed from Spanish.

Table 29. Some Day Names Borrowed from French.

Haiti Creole	**Michif** (N.Dakota)	**Carrier** (Central BC)	**Esperanto** (Invented 1887)
Dimanche	Jimawnsh	Dimosdzin (dzin=day)	Dimanĉo
Lindi	Laenjee	Landi	Lundo
Madi	Marjee	Whulhnatdzin (a) (nat=2)	Mardo
Mécrédi	Mikarjee	Whulhtatdzin (a) (tat=3)	Mercredo
Jodi	Zhweejee	Whulditdzin (a) (dit=4)	Jaudo
Vénnerédi	Vawndarjee	(b)	Vendredo
Sâmedi	Samjee	Sumdi	Sabado (a)

(a) Name not borrowed from French. (b) Not available.

Summary

The names of the days of the week follow either the planetary convention (e.g. in Dutch, Hindi, Japanese, or Quechua) or the numeric convention (e.g. in Portuguese, Russian, Mandarin, or Swahili). These two conventions are well over 2000 years old and between them account for most of the languages spoken today. Religion-related day names, such as *Ijumaa* in Swahili, or *Voskresyénye* in Russian, are also common, but are usually limited to one or two days of special significance, most commonly Friday, Saturday, or Sunday. In most languages the forms of the day names have changed since their initial introduction, so that today's speakers are typically unaware of their origin.

Acknowledgment

The author wishes to express his gratitude to informants and linguists who supplied him with information and advice on many languages.

REFERENCES

AVENI, ANTHONY. 1995. *Empires of time*. New York: Kodansha, America, Inc.

BAZIN, LOUIS. 1989. Les noms des jours de la semaine planetaire dans les langues turques, *Studia Iranica*—Cahier 7— 23-36, Paris, France .

DARDEL, ROBERT DE. 1996. Les noms des jours de la semaine en protoroman. *Revue de Linguistique Romane*, Nancy, France, 60: 321-334

DUNCAN, D.E. 1998. *Calendar*, Avon Publ.

ENCYCLOPEDIA BRITANICA 1983. Fifteenth Edition

HOLMAN, ROBYN A. 1994. The Romance days of the week: An underlying cohesiveness. *The Language Quarterly*, Tampa: 32: 3-4 165-174.

KRUEGER, JOHN R. 1962. Mongolian Personal Names. *Names. Journal of the American Names Society*. 10:81-86.

MOPOHO, RAYMOND. 1998. Personal communication.

O'CROININ, DAIBHI. 1981. The oldest Irish names for the days of the week. *Eriu*. 32: 95-114.

O'NEIL, WILLIAM M. 1978. *Time and the calendars*. Sydney: Sydney University Press.

SALZMAN, MICHELE R. 1990. *On Roman time. The codex-calendar of 354*. Berkeley: Univ. Calif. Press.

STROSS, BRIAN. 1983. Oppositional pairing in Mesoamerican divinatory day names. *Anthropological Linguistics* 25: (2), 211-273.

TRASK, LARRY. 1998. "Days and months in Euskara". Buber's Basque page on the Internet http://weber.u.washington.edu/~buber/Basque/Euskara/days

WILSON, P.W. 1937. Chapter 9. *The romance of the calendar*, New York: W.W. Norton.

AP CLEFTABILITY

Cheng Luo
Brock University

ABSTRACT

Clefting as a syntactic process is widely applicable to NPs and PPs but not to APs and VPs. However, in certain discourse contexts and structural configurations, APs can become more cleftable. Of special interest are colour terms and predicational APs, whose cleftability varies in different contexts. With regard to colour terms, Heggie's (1988) nominal analysis fails to capture neither the distinct adjectival features of colour terms nor those colour terms that are morphologically marked with overt adjectival affixes. Alternatively, this paper proposes a nominality analysis that examines the extent to which colour terms and other APs exhibit nominal features rather than treating them as categorically nominal. A generalized Nominality Principle is then proposed, and predicational APs further examined in terms of cleftability and contextually generated specificity. It is found that predicational APs become more cleftable if (i) the predicate has a restrictive modifier, which provides the cleft focus with an exclusive meaning; (ii) the cleft focus has a strong, contextually generated contrastive reading; and/or (iii) the clefted AP exhibits nominal properties structurally. Finally, AP cleftability is also found to be constrained by interaction between specificity and syntactic complexity.

1 AP Cleftability and the Nominality Principle

Clefting as a syntactic process is widely applicable to NPs and PPs but not to APs and VPs, due to the latter's non-argument status (Emonds 1976). For example, clefting of APs in English results in ungrammaticality, as in (1):

(1) a. *It is quiet that the e man lost his temper. (Heggie 1988:227)
b. *It is happy that he is e.
c. *It's mellow that Bill [finds] Susan e.

In terms of grammatical functions, the clefted AP is attributive in (1a), predicative in (1b), and the predicate of a small clause in (1c). In all the cases, the APs are uncleftable. Another category which is treated as AP in this study is predicational NPs of indefinite reference, which are usually felt to be adjectival in nature, even if they take the form of NPs. This is supported by the fact that, like adjectives, such predicational NPs, when relativized, take the COMP *which* rather than *who* in non-restrictive clauses with a [+human] antecedent, as in (2):

(2) a. This is Mary, who/*which you have already met.
b. She is clever, which/*who you are not.
c. She is *a genius*, which/*who you are not.

In (2c), the relative pronoun of the predicational NP antecedent *a genius* follows an adjectival pattern as seen in (2b), rather than a nominal pattern as in (2a). Moreover, such predicational NPs are in general as uncleftable as predicative APs, as in (3):

(3) *It is a genius that she is e. (cf. (1b))

Structural properties like these, then, justify treating such predicational NPs as APs.

However, in certain discourse contexts and structural configurations, APs can become more cleftable (Declerck 1983, 1984b; Quirk et al. 1985; Pinkham & Hankamer 1975; Lumsden & Lefbvre 1990). One kind of apparent counterexamples, as noted in Quirk et al. (1985), Heggie (1988), and Pinkham & Hankamer (1975), are colour adjectives used predicatively, as shown in (4):

(4) It's **dark green** that we've painted the kitchen e. (Quirk et al. 1985)

Several possible explanations are discussed in Heggie (1988) with regard to cleftability of colour adjectives as shown in (4). The first one simply treats colour adjectives as an *ad hoc* class of adjectives which are somehow susceptible to clefting. However, it falls apart if only one considers examples like (5), where colour adjectives used predicatively are as uncleftable as other APs:

(5) a. *It's blue that the barn could be e. (Heggie 1988:228)
b. *It's white that I saw John turn e.

The second possible explanation takes into account the semantic aspect of the clefted colour adjective and attributes its cleftability to its resultative interpretation in sentences like (4). This is again quickly dismissed by examples like (6) which show that resultative APs are in general uncleftable:

(6) a. *It's tough that Bill cooked the meal e. (Heggie 1988:28)
b. *It's flat that they hammer the nail e.

A third explanation, suggested by Heggie (1988), analyzes the clefted colour term in (4) as an optional argument of the verb *paint*, i.e. the colour term is in fact nominal rather than adjectival. Evidence in support of this analysis includes first of all semantic relatedness between the verb and the colour term. More specifically, while a colour is implied with a verb such as *paint*, 'there is no similar implication of "toughness" with the verb *to cook* in (6a) or of "flatness" with the verb *to hammer'* in (6) (Heggie 1988:229). Moreover, a comparison between the interrogative WH-words used in WH-questions based on (4) and (6b) lends further support to the nominal analysis of colour terms: while *what*, a typical argument operator, is used to replace the colour term in (7a), *how*, an adjunct operator, is used in place of the adjective in (7b) of a similar structure:

(7) a. A: What/*How did they paint the house?
B: Dark green.

b. A: How/*What did they hammer the nail?
B: Flat.

Thus the colour term in (4) should be analyzed as an NP argument and, as such, is cleftable like other NPs.

Heggie's analysis of clefted colour terms not only accounts for the morphological difference

revealed in (7) as well as cleftability of colour adjectives, it is, in my view, also suggestive of a more general principle regarding cleftability, *viz.* the nominal properties of cleftable elements, especially cleftable non-NP constituents.

Despite its plausibility, Heggie's analysis, as I will show, fails to capture neither the distinct adjectival features of colour terms nor those colour terms that are morphologically marked with overt adjectival affixes. I contend that examples like (4) should be more plausibly analyzed as an adjective having acquired some nominal features and become more noun-like. The argument proceeds as follows. Consider Example (8):

(8) It's **greenish** that they painted the house e.

Unlike (4), in which the colour term is morphologically unmarked and can be interpreted as either adjectival or nominal, in (8), the adjectivizing suffix *-ish* unequivocally marks *greenish* as an adjective, which, like the clefted colour term in (4), can be questioned with *what* rather than *how*:

(9) A: What/*How did they paint the house?
B: Greenish.

Given that *greenish* can only be adjectival, Heggie's explanation would not apply here to account for grammaticality of (8). On the other hand, *greenish* and other colour terms do show the nominal property of being substitutable by *what* but not by *how* in WH-questions. The question remains whether there is a more principled and plausible way of accounting for cleftability of colour terms in particular and AP cleftability in general.

What I suggest is a nominality analysis from a prototypical perspective, which differs from Heggie's nominal analysis by examining the extent to which colour terms and other APs exhibit nominal features rather than treating them as categorically nominal; and it is this increased degree of nominality, together with semantic relatedness noted in Heggie (1988), that makes colour terms more cleftable. Therefore, treating colour terms not as inherently nominal but as adjectives with certain degree of nominality enables us to account for both increased cleftability of colour terms in given contexts and their inherent adjectival features such as gradability.

As suggested above, the nominal properties of colour terms in relation to their cleftability may be generalizable to a principle regarding cleftability, especially that of non-NP constituents. On the basis of the nominality analysis, and following Pinkham & Hankamer (1975), I will propose a 'Nominality Principle' to account for cleftability of non-NP constituents in general that also subsumes NP cleftability:

(10) **The Nominality Principle**
The more noun-like a constituent is, the more cleftable it tends to be

where being more noun-like is defined both structurally and semantico-pragmatically. Structurally, a non-NP is more noun-like if it shows morphological and/or syntactic markings that are usually associated with NPs. Pragmatically, a non-NP is more noun-like if it exhibits semantico-pragmatic features typically associated with NPs, such as specificity and definiteness. Let's now turn to the notion of specificity.

2 Cleftability of predicational APs: the Specificity Condition

As noted above, a cleft focus that functions as an underlying post-copula predicative in the presupposed clause is unacceptable:

(11) a. *It is a genius that he is e. (Leech & Svartvik 1975:181)
b. *It is the football coach that John is e. (Emonds 1976:140)

Based on such observations, Emonds (1976:140) makes the assertion that 'predicative nominatives and predicate adjectives do not appear in focus position in the cleft construction.' Lee (1963:380) had expressed a similar view. However, as will be shown, this asserted restriction on cleftability of predicative elements can be reversed or relaxed depending on context and according to a semantico-pragmatic condition which conforms to (10).

In Akmajian (1979) and Higgins (1976), copula sentences are classified into 'specificational' and 'predicational'. A copular sentence is specificational if one of the NPs (which is usually the subject NP) represents a variable for which the other NP (which is usually the predicative NP) specifies a value:

(12) A: Who's the chairman of the committee?
B: The chairman is John Smith.

In (12), the NP representing the variable (i.e. *chairman*) normally resembles the heading of a list, the value of which is specified by the predicative definite NP *John Smith*. On the other hand, a copular sentence is predicational if the predicative NP/AP does not represent a value specified for a variable but functions as a sort of semantic predicate, in the sense that it does not identify the referent of the subject NP but simply provides more information about it. Usually, the predicational NP in this case describes a property or a role, or indicates class membership:

(13) a. Mary is a pretty girl/a teacher.
b. He is a Canadian.

For example, in (13a), *a pretty girl* describes a property of the subject NP, *Mary*; and *a teacher* states a role of *Mary*. In (13b), the predicational NP describes the class membership (i.e. *Canadian*) of the subject NP. Unlike the subject NP in a specificational sentence which represents a variable, the subject NP in a predicational sentence is already quite definite. Functionally, the post-copula element is identificational in a specificational sentence but descriptive in a predicational one. Structurally, while the pre-copula and post-copula NPs in specificational sentences are reversible, those in predicational sentences are not, as in the ungrammatical **A teacher is Mary*.

The cleft sentence is essentially a specificational type of sentence, in that the cleft focus is almost always identificational and the presupposed part always represents a variable. An important feature required of the value assigned to a variable is that it needs to be specific by virtue of having some exclusive meaning. It is this specificity (and related exclusiveness) that affects cleftability of predicative elements in predicational sentences. As attested earlier, such predicative elements, even if they may take the form of an NP, are usually felt to be adjectival; and as such are not felt to be exclusive in meaning, because one characteristic does not automatically exclude other characteristics

someone may have. For example, as Declerck (1984) argues, to say *Mary is pretty* does not exclude diligence as another characteristic that Mary may have. If we relate such lack of exclusiveness to low specificity, we may say that a predicational element is generally uncleftable because it fails to meet the specificity condition of a cleftable constituent.

On the other hand, where we do find cleft sentences with a predicational element as the cleft focus, we also tend to find increased exclusiveness, and subsequently increased specificity, in the predicational element, as in (14) and (15):

(14) A: What is Mary's most typical characteristic? (Declerck 1984:135)
B: It is pretty that Mary is e, more than anything else.

(15) a. *It is a good citizen that he is not.
b. If there is one thing that he is not, it is a good citizen.

The increased exclusiveness in the predicational elements in (14) and (15) is generated by the context. In (14), the answer in the context of the preceding question implies that being pretty and not any other characteristic is most typical of Mary. Similarly, while (15a) is unacceptable if used in isolation, (15b) is much better because the nominal idea (i.e. *the thing that he is not*) and exclusiveness are contextually generated. Therefore, for a predicational element to be the cleft focus, it needs to be specific in the context. This can be stated as the Specificity Condition in (16):

(16) The Specificity Condition
An predicational element is not cleftable unless it is made more contrastively specific by the context.

3 Contextually generated specificity

One way specificity can be increased is through use of modifiers such as various kinds of adjuncts, which specify the scope within which the proposition is true. Generally, modified predicational elements are more specific, and more cleftable, than bare predicational elements. To wit,

(17) a. *It is happy that he is e.
b. ?It is happy that he always is e whenever I see him.

(18) a. *It was willing that he appeared e.
b. †It is willing that he must appear e, if he is to succeed in this interview.

The four sentences in (17) and (18) differ in two ways: (i) whether the predication is modified by some adjunct, and (ii) whether the link verb is *be*. We note that while (17a) and (18a) have unmodified predicates, (17b) and (18b) both have adjuncts which create a rather specific setting for the predicates. On the other hand, whereas (17) use the copula *be*, (18) use the link verb *appear*. Semantically, non-*be* link verbs are less depleted, i.e. more substantial, than the copula *be*, and therefore more specific. As Declerck (1984:145) points out, unlike *be*, link verbs like *look, become, appear, grow*, etc. 'express a predicational relation that is more readily linked up with specific

circumstances or with a specific time.' The property assigned to the referent of the subject NP is therefore more easily felt to be exclusive: it is the particular property that the referent has or had in specific circumstances. Therefore, both presence of adjuncts and the use of a non-*be* link verb increase specificity for the clefted AP in (18b), making it the most cleftable of the four sentences.

Another way to examine specificity of a clefted predicational element is to see whether it has a strong contrastive reading from the context. Contextually, it is possible to identify cleft sentences as having a weak contrastive reading or a strong contrastive reading, as illustrated in (19):

(19) a. A: Who hit John? (Heggie 1988:205)
B: It's Mary who e hit John.

b. A: Who hit John?
B: Mary did.
C: No, it's Bill (not Mary) who e hit John.

The difference between (19a) and (19b) is that while (19a) only gets a weak contrastive reading, (19b) has a strong contrastive reading in relation to a prior utterance. Despite this difference, both (19a) and (19b) are grammatical.

The significance of the contrast between weak and strong contrastive reading is more clearly seen in clefted predicational APs, as shown in (20).

(20) a. A: What colour are her eyes? (Heggie 1988:206)
B: *It's green that her eyes are e.

b. A: What colour are her eyes?
B: Her eyes are green.
C: No, it's BLUE that her eyes are e, not GREEN.

(20) parallels (19) in that contextually, (20a) gets a weak contrastive reading whereas (20b) a strong contrastive reading. It differs from (19) in that while the degree of contrastiveness does not affect grammaticality of (19), where NPs are clefted, it does negatively affect cleftability of predicational APs with a weak contrastive reading, as shown in (20a). More importantly, as (20b) shows, predicational APs tend to become more cleftable when made more contrastively specific by the context. Also, NPs are more cleftable than APs by virtue of wider distribution (e.g. in clefts with both strong and weak contrastive reading). In other words, the less nominal constituent is syntactically more constrained with regard to clefting. As another example of greater cleftability being reflected in wider distribution, consider (21):

(21) a. Susan said that it was John that e Mary hit and Jane said it was Tom.
b. *Susan said that it's stupid that John is e and Mary said that it's clever.

Both (21a) and (21b) have coordinate clauses embedded in the matrix sentence, with the VP deleted from the second of the conjoined clauses under identity. However, while an NP can be clefted from such an embedded clause, as in (21a), a predicational AP cannot, as in (21b). Interestingly, a comparison between (20b) and (21b) reveals further that the Specificity Condition alone may not

fully account for cleftability of predicational APs; we have to consider syntactic complexity such as embedding, such that the more deeply embedded a constituent, the less cleftable it tends to be. More specifically, in (21b) the context yields a strong contrastive reading similar to that in (20b), but since the clefting in (21b) is from embedded clauses whereas that in (20b) is from a matrix sentence, (20b) allows clefted APs but (21b) does not. This is due not to specificity, but to syntactic complexity: even if a predicational AP is made more specific by the context, syntactic constraints may still prevent it from being clefted. Therefore, both semantico-pragmatic factors and syntactic factors, and their interaction thereof, should be considered when dealing with AP cleftability.

4 Conclusion

Generally speaking, APs are either uncleftable or have very low cleftability compared with NPs. However, through a nominality analysis, this paper has shown that predicational APs may become more cleftable under the following conditions:

(22) a. the predicate has a restrictive modifier, which provides the cleft focus with an exclusive meaning that in turn makes the AP more specific;

b. the cleft focus has a strong, contextually-generated contrastive reading which makes an AP more specific; and/or

c. the clefted AP exhibits nominal properties structurally.

Finally, AP cleftability is also found to be constrained by interaction between specificity and syntactic complexity.

REFERENCES

Atlas, J. & Levinson, S. 1981. 'It-clefts, informativeness, and logical form: radical pragmatics,' (revised standard version). In P. Cole (ed.). *Radical Pragmatics*. New York: Academic, 1-61.

Declerck, R. 1983. 'Predicational clefts,' *Lingua, 61, 1*: 9-45.

--------. 1984a. 'The pragmatics of *it*-clefts and wh-clefts,' *Lingua, 64, 4*: 251-89.

--------. 1984b. 'Some restrictions on clefts that highlight predicate nominals,' *Journal of Linguistics, 20, 1*: 131-154.

Delahunty, G. 1984. 'The analysis of English cleft sentences,' *Linguistic Analysis, 13, 2*: 63-113.

Delin, J, & Oberlander, J. 1995. 'Syntactic constraints on discourse structure: the case of *it*-clefts,' *Linguistics: An Interdisciplinary Journal of the Language Sciences, 33, 3*: 465-500.

Emonds, J. 1976. *A Transformational Approach to English Syntax*. New York: Academic.

Fichtner, E. 1993. 'Cleft sentences in English: a comprehensive view,' *Word 44, 1*: 1-30.

Geluykens, R. 1991. 'Discourse functions of *it*-clefts in English conversation,' *Communication and Cognition, 24, 3-4*: 343-57.

Heggie, L. 1988. *The Syntax of Copular Sentences*. Doctoral dissertation, Los Angeles: University of Southern California.

Horn, L. 1981. 'Exhaustiveness and the semantics of clefts.' In V. Burke & J. Pustejovsky (eds.). *Proceedings of the Eleventh Annual Meeting of the North Eastern Linguistic Society.* Amherst, MA: Graduate Ling. Student Assn., Univ. of Massachusetts, 125-142.

Kamio, A. 1991. 'Cleft sentences and the territory of information.' In C. Georgopolous. & R. Ishihara (eds.). *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda.* Dordrecht: Kluwer, 353-71.

Lefebvre, C. 1992. 'Towards a typology of predicate cleft languages,' *Journal of West African Languages, 22, 1*: 53-61.

Lumsden, J. & Lefebvre, C. 1990. 'Predicate cleft constructions and why they aren't what you might think,' *Linguistics: 28,* 4: 761-782.

Luo, C. 1994. 'Thematicity and cleftability.' In D. Jory & V. Motapanyane (eds.). *Papers from the Eighteenth Annual Meeting of the Atlantic Provinces Linguistic Association.* University of New Brunswick, Saint John, N.B.: Atlantic Provinces Linguistic Association, 54-73.

Pinkham, J. & Hankamer, J. 1975. 'Deep and shallow clefts,' *Papers from the 11th Regional Meeting of the Chicago Linguistic Society.* Chicago: CLS, University of Chicago, 429-50.

Quirk, R. et al. 1985. *A Comprehensive Grammar of the English Language.* London: Longman.

Vion, M. & Colas, A. 1995. 'Contrastive marking in French dialogue: why and how'. *Journal of Psycholinguistic Research, 24, 5*: 313-31.

Syntaxe standard ou non standard: remplacements et contournements des relatives chez l'enfant acadien

Agnès Picolet-Crépault
Université Sainte Anne

RÉSUMÉ

La présente étude porte sur les remplacements et les contournements en français parlé acadien des pronoms relatifs objet du Sud-Est du Nouveau-Brunswick. A l'aide de deux épreuves, l'une de compréhension, l'autre de production dirigée, nous analysons les résultats obtenus. Nous constatons que la syntaxe de l'enfant acadien de six à douze ans comporte une forte proportion de contournements dans l'épreuve de production dirigée quelque soit l'année de scolarité. Les relatives sont contournées soit par des énoncés simples coordonnés ou juxtaposés, soit par des relatives sujet comportant souvent des constructions complexes du type passif-pronominal. Nous en concluons que la relative objet n'est pas une structure facilement accessible chez l'enfant acadien en milieu minoritaire car deux sujets (de la principale, de la subordonnée) entrent en concurrence.

La langue parlée des Acadiens procède de l'héritage linguistique historiquement légué aux francophones de cette région, immergés dès l'enfance dans un double sytème d'expression, le français et l'anglais et exposés à d'autres variétés de langue, en particulier le français familier ou, dans la région de Moncton, le << chiac >> , sorte de << métissage >> linguistique (1).

Au carrefour du système standard et non standard, les relatives du Sud-Est du Nouveau-Brunswick comportent certaines spécificités, notamment l'emploi de structures décumulées fréquentes du type *que... de* ou *que... à* (Picolet-Crépault, 1997). En outre, nos recherches indiquent qu'une analyse pointue des contournements serait susceptible de contribuer de manière significative aux travaux en cours. Les termes << remplacements >> ou << contournements >>, fréquemment employés tant par les linguistes (Dubuisson, C. et Ermikanian, L., 1982, Gadet, F., 1995) que par les psycholinguistes (Deyts et Noizet, 1973), désignent la capacité du locuteur à remplacer ou à contourner une structure par une autre. Nous limiterons nos données aux relatives objet.

(1) Les Acadiens sont les descendants de colons français de la Nouvelle-France, dont les ancêtres étaient en majorité originaires du Poitou-Charente. Ils se distinguent des Québécois par l'histoire et des particularités linguistiques. Malgré l'ostracisme dont ils furent victimes après la conquête des Anglais, notamment leur déportation en 1755 dans les colonies anglaises d'Amérique, ils ont réussi à conserver leur langue et leur coutumes en dépit d'une inquiétante assimilation.

Aujourd'hui, ils sont disséminés dans les provinces du Nouveau-Brunswick, de Nouvelle-Écosse et de l'Ile-du - Prince-Edouard où survivent des régions à prédominance acadienne. Le recensement de 1996 montre que la population du Nouveau-Brunswick comprend 33,2% d'Acadiens de langue maternelle française dont 30,5% ont le français comme langue d'usage. Depuis 1967, les droits linguistiques des francophones au Nouveau-Brunswick sont officiellement reconnus par une loi officialisant le bilinguisme dans la province.

Méthode expérimentale

L'administration de deux épreuves consécutives - compréhension et production dirigée - nous permettra de déterminer la méthode la plus appropriée pour découvrir les stratégies syntaxiques à l'oeuvre chez le jeune locuteur. Notre analyse sera distributionnelle.

La population est composée de quatre-vingt-dix enfants acadiens âgés de six à douze ans de catégories socio-culturelles moyennes de la région de Moncton. Elle est divisée en six groupes de quinze élèves en proportion égale des deux sexes sur six années consécutives de scolarité, de la 1ère à la 6ème année.

Les épreuves sont administrées individuellement à chaque enfant. La langue parlée au foyer est française et avec quelques incidences d'anglais pour 78% des pères (2) et 87% des mères. Dans une étude précédente (Picolet-Crépault, 1993), nous avions noté la plus grande fréquence d'emploi du métissage français-anglais ou de l'anglais par les pères que par les mères.

Comprehension

Méthode

L'expérimentateur commence par donner verbalement un bref énoncé descriptif et invite ensuite l'enfant auquel il s'adresse à désigner dans un jeu de douze images la phrase qui correspond à l'énoncé qu'il vient d'entendre. Les énoncés, présentés en ordre aléatoire, comportent des propositions relatives sujet, objet et à cas oblique (prép+lequel). Si les jeunes sujets réussissent en général bien l'épreuve, on constate néanmoins une moindre maîtrise des relatives objet en *que* et à cas obliques comparativement aux relatives sujet en *qui*.

Les propositions relatives objet standard sont les suivantes (3) :

(1) Premier récit:
C2: *L'oiseau | que la fille caresse | mange des graines*
C4: *La petite fille s'éloigne de la cage | qu'elle a laissée ouverte*

(2) Deuxième récit:
C7: *Le chien | que le garçon abandonne | reste seul*
C9: *Le chien suit la piste de l'ami | qu'il a perdu*

---

(2) Influence du milieu de travail: peut être causé par la situation fréquente des travailleurs acadiens encadrés par des agents de maîtrise unilingues anglophones.

(3) Dans les phrases C2 et C7 ci-dessus, la relative est dite *enchâssée*, c'est à dire divisant la proposition principale en deux segments tandis que dans les phrases C4 et C9, elle est dite *juxtaposée*, la proposition relative suivant la proposition principale.

Résultats

Tableau 1

Épreuve de compréhension : Fréquence des remplacements

Années Scolarité	1e	2e	3e	4e	5e	6e	Total occur.
Phrases: C2	1	3	3	2	1	0	10
C7	5	1	2	2	1	0	11
C4	3	1	0	0	0	0	4
C9	1	0	0	0	0	0	1
Total /360	10	5	5	4	2	0	26

L'examen de ces résultats selon une approche distributionaliste montre que:
-les relatives objet en *que* (surtout les enchâssées) suscitent davantage de remplacements que les relatives sujet en *qui;*
-les enchâssées en *que* présentent plus de difficultés que les juxtaposées;
-la fréquence de remplacements diminue au cours des trois dernières années avec la progression de l'élève dans sa scolarité (de 20 à 6 occurrences).
Les remplacements portant sur les propositions relatives objet sont réalisés de deux manières:

1) par une relative enchâssée en *qui*
Par exemple, la phrase:
C2: *L'oiseau | que la fille caresse | mange des graines*
est remplacée par la phrase:
C3: *La fille | qui ouvre la porte | met des graines dans la mangeoire*

2) par une relative juxtaposée en *qui*
Ainsi, la phrase:
C7: *Le chien | que le garçon abandonne| reste seul*
est remplacée par la phrase:
C10: *Le chien regarde par la fenêtre les enfants | qui travaillent |*

Analyse

Ces résultats corroborent les nombreuses études publiées sur les difficultés des relatives en *que*, en particulier l'enchâssée (Slobin, 1973, Sheldon, 1974, De Villiers et al., 1979). La production de l'enchâssée en *que* présentant des fonctions non parallèles (objet-sujet), exige de l'enfant un processus complexe de recherche de l'antécédent, qui le conduit à interpréter le relatif *que* comme le relatif *qui,* l'antécédent gardant pour le jeune locuteur une fonction sujet dans les deux cas.

Production dirigée

Méthode

L'expérimentateur propose à l'enfant de consulter douze images différentes l'une après l'autre présentées en ordre fixe et de produire sur chaque image une phrase comportant une relative objet.

-L'expérimentateur pointe du doigt l'objet de départ.
-L'enfant doit trouver la phrase correspondant à l'image.
-Si le sujet hésite, l'expérimentateur lui repète plus clairement la consigne.

Les réponses attendues peuvent être les suivantes:

PD5:*C'est l'eau | que le chien boit |*
PD6: *C'est la balle | que le chien attrape*
PD7: *C'est le livre | que la fille lit |*
PD8: *C'est le ballon | que la fille lance |*

Résultats

Tableau 2

Épreuve de production dirigée: fréquence des contournements

Années scolarité	1e	2e	3e	4e	5e	6e	Total /360
Phrases PD5	8	10	9	8	7	8	50
PD6	8	13	11	7	5	6	50
PD7	4	13	5	5	1	9	37
PD8	9	10	8	4	3	6	40
Total	29	46	33	24	16	29	177

Nous constatons :

-une légère diminution des contournements proportionnelle à l'avancement dans la scolarité dans les trois dernières années (108 contre 69 occurrences);
-une plus forte proportion des contournements comparativement à l'épreuve de compréhension (179 contre 26 occurrences en compréhension);
- des contournements de deux types, par énoncés simples et par relatives sujet en *qui.*

Contournements par énoncés simples.

L'épreuve a donné lieu à 89 occurrences.

Tableau 3

Fréquence des contournements par années de scolarité

Années	1e	2e	3e	4e	5e	6e
ES /360 occ.	21	31	11	9	7	10

L'énoncé simple peut comporter des structures récurrentes:
- comportant un adjectif
  1. *C'est un chien noir et blanc*
- une cible génitive
  2. *C'est la balle au chien*
- une cible circonstancielle (de lieu ou d'accompagnement).
  3. *C'est de l'eau dans un bol*
  4. *C'est un livre avec des paroles écrites dedans*

-coordonnées ou juxtaposées :
  5. *C'est un livre, la petite fille est en train de lire le livre*
  6. *C'est l'eau pi le chien la boit*

On constate que les contournements par énoncés simples régressent sensiblement à partir de la troisième année de scolarité (26 occurrences contre 63 occurrences dans les trois premières années).

Contournements par les relatives sujet en *qui*.

L'épreuve a donné lieu à 88 occurrences.

Tableau 4

Fréquence des contournements en *qui* en fonction de l'âge

Années	1e	2e	3e	4e	5e	6e
/360 occ.	8	15	22	15	9	19

L'examen des fréquences de contournements en *qui* nous indique que cette stratégie est employée quel que soit l'année de scolarité. A première vue, les contournements en *qui* semblent être un procédé de simplification. L'antécédent étant le sujet de la proposition, il n'y a pas d'inversion d'actants et pas de recherche de verbe. Procédons à une analyse qualitative.

Les contournements en *qui* sont de trois types, variant selon la nature du verbe employé. Ils peuvent comporter un verbe:

- actif:

7. *C'est le ballon qui vole dans les airs*

- auxiliaire:

8.*C'est la balle qui est par terre*

- ou passif-pronominal (27 occurrences)

9. *C'est une balle qui va s'faire prendre par le chien*

10. *C'est une balle qui s'fait tirer en l'air*

La fréquence d'emploi d'une proposition relative objet

-comportant un verbe actif est constante à tous les âges (24 occurrences de la 1ère à la 3ème année de scolarité, 12 occurrences de la 4ème à la 6ème année);

-comportant le verbe auxiliaire diminue au cours des trois dernières années (18 occurrences de la 1ère à la 3ème année, 7 occurrences de la 4ème à la sixième année);

-comportant un verbe passif-pronominal augmente avec l'année de la scolarité dans cet échantillon.

Tableau 5

Fréquence des contournements des relatives objet par une relative sujet avec verbe passif-pronominal

Année	1	2	3	4	5	6
Verbe PP	0	2	1	8	4	12

Analyse

D'un point de vue descriptif syntaxique, la structure *que+ se faire* avec ou sans agent comporte, selon Blanche-Benveniste (1984:137), <<un auxiliaire de reformulation dont l'effet de sens est équivalent à certain type participatif>>. En outre, selon Le Goffic (1993:226), <<le verbe *se faire* concurrence le passif... c'est un verbe pratique car le patient n'a pas le rôle passif>>.

D'un point de vue développemental, l'appel à une structure passive-pronominale semble caractériser une étape de progression morphosyntaxique chez l'enfant à partir de dix ans. Deyts et Noizet (1973 : 210) remarquent sur un autre échantillon qu'à partir de douze ans, <<la complication apportée par l'ajout d'une transformation passive est largement compensée par le gain obtenu sur la transformation relative>>. Les enfants semblent donc préférer, dans certains cas, un contournement avec une relative sujet comportant un verbe passif-pronominal à une relative objet impliquant une inversion du sujet.

Ainsi, les contournements ou remplacements employés diffèrent par leur forme et sont de deux types: énoncés simples et relatives sujet. Les enfants acadiens emploient des contournements en *qui* du type passif-pronominal quand ils font face à des contraintes de trois types: limitation des verbes à utiliser, directivité de l'épreuve et enfin, complexité de l'utilisation des relatives objet. Il est intéressant de noter que les relatives passives-pronominales comportent

souvent des formes verbales complexes (passé ou au futur). Ce procédé contredit donc le principe de simplification appliqué aux contournements et aux remplacements syntaxiques.

Comme le soulignait Ferreiro et al. (1976), l'enfant de douze ans n'a pas totalement maîtrisé le système des relatives objet et préfère utiliser soit des structures plus simples soit des structures assez complexes mettant en jeu une syntaxe non maîtrisée.

Conclusion

Une comparaison des résultats respectifs des deux épreuves semble indiquer la moindre pertinence d'une épreuve de compréhension en termes de connaissance de la syntaxe et des stratégies de l'enfant qu'une épreuve de production dirigée. Pour le jeune locuteur, il ne s'agit plus d'un jeu de devinettes entre images et phrases d'adultes mais bien d'un exercice complexe révélant ses operations intellectuelles.

L'épreuve de production dirigée permet de mettre en évidence une forte fréquence de contournements des relatives standard objet chez l'enfant acadien de six à douze ans en comparaison avec les enfants français du même groupe d'âge (Picolet-Crépault, 1998). Une approche pédagogique qui vise à enseigner le français standard sans imposer ni rejeter en acceptant le français parlé régional acadien et les multiples contournements du français standard amène l'élève à un enrichissement de son potentiel linguistique sans abandonner sa langue première.

BIBLIOGRAPHIE

BLANCHE-BENVENISTE, C., DELOFEU, J., STEFANINI, J. et VAN DEN EYNDE, K. 1984. *Pronom et syntaxe. L'approche pronominale et son application en français.* Paris: SELAF.

DEYTS, J.P. et NOIZET, G. 1973. <<Étude génétique de la production de la subordonnée relative>>. *Cahiers de Psychologie*, vol.16, pp. 199-212.

DUBUISSON, C., ERMIKANIAN, L. 1982. <<Acquisition des relatives et implications pédagogique>>. Dans C. Lefebvre. *La syntaxe comparée du Français standard et populaire*, vol.1, UQUAM, Montréal, Québec, pp. 367-391.

FERREIRO, E., OTHENIN-GIRARD, C., CHIPMAN, C. et SINCLAIR, H. 1976. <<How Do Children Handle Relative Clauses?>> *Archives de Psychologie*, vol.44, 172, pp. 229-266.

GADET, F. 1995. <<Les relatives non standard en français parlé: le système et l'usage>>. Copenhague, *Études Romanes*, vol.34, pp. 141-162.

LE GOFFIC, P. 1993. *Grammaire de la phrase française.* Paris: Hachette Supérieur.

PICOLET-CRÉPAULT, A. 1993. <<Bilinguisme additif ou soustractif chez les enfants de 6 ans dans une école du Sud-Est du Nouveau-Brunswick >>. Actes des 7e journées de linguistique, *Revue de l'AEDILL*, CIRAL, Québec, Université Laval.

PICOLET-CRÉPAULT, A. 1997. <<Étude du langage enfantin: les propositions relatives chez l'enfant acadien de six à douze ans>>. *Thèse de Doctorat* (Ph.D). Université de Moncton, Canada.

PICOLET-CRÉPAULT, A. 1998. <<Étude comparative des propositions relatives chez les enfants d'Aix-en-Provence>>. *Rapport Post-Doctoral*. Université d'Aix-Marseille, France.

SHELDON, A. 1974. <<The Role of Parallel Function in the Acquisition of Relative Clauses in English>>. *Journal of Verbal Learning and Verbal Behaviour*, vol.13, pp. 305-318.

SLOBIN, D.I. 1973. <<Cognitive Prerequisites for the Development of Grammar Studies>>. Dans C. Ferguson. et D.I. Slobin. *Studies in Child Language Development*. New York: Holt, Rinehart and Winston, pp. 175-208.

VILLIERS (DE), J., FLUSBERG, H., HAKUTA, K. et COHEN, M. 1979. <<Children Comprehension of Relative Clauses>>. *Journal of Psycholinguistic Research*, vol.8, pp. 400-517.