



LES AUTOCORRECTIONS IMMÉDIATES DANS LA PAROLE LUE EN KIRUNDI : FRÉQUENCE ET MÉCANISMES DE RÉPARATION

Gélase NIMBONA

Centre de recherche en Langues, Cultures et Sociétés (CRELACS), Département des Langues et Cultures Africaines, Faculté des Lettres et Sciences Humaines, Université du Burundi (Bujumbura, Burundi)

Mail : gelase.nimbona@gmail.com

Reçu : 08/10/2025 ; Accepté : 29/11/2025 ; Publié : 31/12/2025

Cite As:

Nimbona, G. (2025). *Les autocorrections immédiates dans la parole lue en Kirundi : fréquence et mécanismes de réparation*. Revue de l'Université du Burundi – Série : Sciences Humaines et Sociales, 22 (2), pp20-30.

Résumé

Contexte. Le kirundi est une langue à tons. Mais contrairement au français par exemple où les signes diacritiques sont spécifiés dans le texte, les tons sont produits mais ne sont pas notés à l'écrit. **Problématique.** Cette situation rend difficile la lecture d'un texte kirundi, le lecteur devant tenir compte du sens et du contexte phrastique global pour savoir quel il faut produire. Cet article traite donc un des phénomènes spécifiques à la langue orale : les autocorrections immédiates. Ces phénomènes – aussi conçus comme des marques de travail de formulation (TdF) – viennent perturber la régularité syntaxique des énoncés et empêchent la formation des constituants de discours traditionnellement connus de la langue écrite. **Objectif.** Ces phénomènes étant conçus comme relevant des mécanismes cognitifs généraux du langage, cet article a pour objectif de non seulement étudier leur attestation dans les productions orales Kirundi, mais aussi de vérifier les hypothèses largement confirmées dans les autres langues dont la typologie diffère de celle du Kirundi. **Méthodologie.** L'analyse porte sur un corpus d'environ 29 minutes de parole lue et de 34 minutes de parole spontanée recueillies auprès de 6 locuteurs kirundophones. **Résultats.** Contrairement aux observations faites en français, les résultats montrent que les autocorrections immédiates sont plutôt fréquentes dans la parole lue que dans la parole spontanée. **Conclusion.** Ce travail révèle donc que, malgré leur caractère universel, la réalisation des marques de TdF peut varier d'une langue à l'autre et d'un style de parole à l'autre.

Mots clés : autocorrection, phénomènes de l'oral, disfluences, marqueurs de travail de formulation.

IMMEDIATE SELF-CORRECTIONS IN READ SPEECH KIRUNDI: FREQUENCY AND REPAIR MECHANISMS

Abstract

Context. Kirundi is a tonal language. But unlike French, for example, where diacritical marks are indicated in the text, tones are produced but not written. **Problematic.** This situation makes reading a Kirundi text challenging, as the reader must rely on meaning and overall sentence context to know which tone to produce. This article addresses therefore of the phenomena specific to spoken language: immediate self-corrections. These phenomena – also conceived as Formulation work marks – disrupt the syntactic regularity of utterances and prevent the formation of discourse constituents traditionally known from written language. **Objective.** Since these phenomena are considered as to be part of the general cognitive mechanisms of language, the aim of this article is not only to study their attestation in Kirundi oral productions, but also to verify hypothesis that have been widely confirmed in other languages whose typology differs from that of Kirundi. **Methodology.** The analysis concerns a corpus of approximately 29 minutes of read speech and 34 minutes of spontaneous speech collected from 6 Kirundi speakers (4 men and 2 women). **Results.** Contrary to the observations made in French, the results show that immediate self-corrections are rather frequent in read speech. **Conclusion.** This work therefore reveals that, despite their universal nature, the realisation of speech repair markers can vary from one language to another and from one speech style to another.

Key words: self-corrections, oral phenomena, disfluencies, formulation work marks.

Introduction

Les autocorrections comme les répétitions sont des marques de formulation plus complexes qui interrompent la linéarité du discours et qui participent à des achoppements propres aux productions orales. En effet, il a été démontré que le caractère oral de la langue parlée induit l'apparition d'inattendus structurels tels les hésitations, les répétitions, les corrections, les inachèvements, etc. (Goulian et Antoine 2001). Et ces inattendus structurels viennent perturber la régularité syntaxique des énoncés, contrariant ainsi la formation des constituants [de discours] traditionnels (Passot 2004). Cela pour ainsi dire que tout énoncé oral spontané conserve les traces de son élaboration à travers des phénomènes de performance, tels que la répétition, l'autocorrection, l'allongement de la finale, etc., qui constituent de précieux indices susceptibles d'éclairer le fonctionnement de la langue (Henry, Campione, et Véronis 2004; Valli et Véronis 1999). Plus spécifiquement, il a été démontré que les pauses silencieuses non structurantes et les autres hésitations sont liées à la difficulté langagière et qu'elles sont les symptômes des difficultés rencontrées dans le traitement et la planification de la parole (Cenoz 1998) ; les pauses (silencieuses et remplies) se révélant comme des indications que le locuteur met du temps à chercher l'élément linguistique qui suit ou à formuler l'idée à venir dans le cours de son discours (Cenoz 1998). Par ailleurs, ces inattendus structurels – qualifiés de marqueurs de travail de formulation (Candea 2000) – ne sont ni aléatoires ni dénuées de fonction. Ils contribuent au processus d'encodage et permettent à l'auditeur le décodage et la compréhension du message oral (voir aussi (Nimbona 2010) pour détails).

Nombre de travaux ont été effectués sur les pauses et les marques de travail de formulations aussi bien dans l'espace anglophone que dans l'espace francophone (Candea 2000, 2001; Dister 2007, 2008; Duez 1999, 2001; Goldman-Eisler 1972; François Grosjean et Deschamps 1975; Levelt 1983; Shriberg 1999; Valli et Véronis 1999). Cependant, nous observons une quasi absence de travaux centrés directement sur l'étude des marques de travail de formulation dans les langues bantu en général et en Kirundi en particulier.

Le présent travail s'intéresse sur les autocorrections immédiates en Kirundi particulièrement dans la parole lue. L'étude part du constat que les phénomènes de performance (répétition, autocorrection, allongement, hésitations vocales) sont des phénomènes caractéristiques de l'oral non contraint et non préparé (Henry 2005), que les pauses sonores n'apparaissent pratiquement pas à l'oral de lecture – l'écrit oralisé (Candea 2000). Notre étude vise à vérifier l'authenticité de cette hypothèse dans l'oral de lecture en Kirundi dont la typologie s'avère différente de celles des langues indoeuropéennes.

Notre objectif est double. D'une part, ce travail cherche à montrer – le psychisme humain étant identique par ailleurs – que les autocorrections immédiates (comme les autres marques de travail de formulation) constituent des indices susceptibles d'éclairer le fonctionnement de la langue quel que soit sa typologie. D'autre part, elle tente de montrer que l'analyse des autocorrections à l'oral de lecture en Kirundi ne peut se faire en se fondant uniquement sur le modèle syntaxique. La lecture du texte Kirundi, plus qu'en français, recèle les processus d'encodage et de décodage et les processus de production et d'encodage se réalisent simultanément. À cette fin, nous avons analysé un corpus de 29 minutes 11 secondes d'enregistrements réalisés auprès de six locuteurs natifs (dont 4 hommes et 2 femmes) en situation de lecture d'un texte journalistique. Ces données de lecture seront analysées en comparaison avec celles de la parole spontanée en vue de saisir la portée des autocorrections en situation de lecture.

Dans la deuxième section, nous éclairons le lecteur sur la terminologie utilisée en vue de situer notre travail dans l'ensemble des autres travaux déjà réalisés sur le même phénomène. La section 3 donne les détails sur les données analysées et la méthodologie suivie, tandis que la section 4 présente les principaux résultats observés dans cette étude avant de passer à la discussion de ceux-ci et à la conclusion dans les sections 5 et 6.

Discussion terminologique

Nous rencontrons dans la littérature une gamme de termes pour désigner la notion d'autocorrection quelle que soit la partie de l'énoncé dans laquelle le locuteur se corrige et indépendamment de la forme que peut prendre cette correction : *faux-départs* (*falsestart*) (F. Grosjean et Deschamps 1972), *révision* (Bénard 2005), *auto-réparation* (de Fornel et Marandin 1996), *repair phenomena* (Schegloff, Jefferson, et Sacks 1977). (Candea 2000) distingue, quant à elle, *autocorrections immédiates* et *faux départs* : les premières sont systématiques alors que les secondes sont plus complexes et concernent aussi des mots pleins ou des structures syntaxiques inachevées. Notre analyse des autocorrections immédiates s'inspire de cette définition que donne Candea. Au niveau pratique, le lecteur remarquera que le fonctionnement de l'autocorrection immédiate telle que nous la concevons dans ce travail est très proche de celui de la répétition. En effet, comme dans le cas des répétitions, la place syntaxique qu'occupe la séquence d'origine selon la terminologie de Levelt (1983) est la même que celle de la séquence d'autocorrection.

Méthodologie

Notre corpus d'étude est constitué des enregistrements effectués auprès des locuteurs kirundophones en situation de lecture et en situation de parole spontanée. Au total, nous avons analysé 12 fichiers dont 6 représentent 29 minutes de corpus de parole lue et 6 autres 34 minutes de parole spontanée. Les six locuteurs (dont 4 hommes et 2 femmes) sont des hommes et des femmes qui sont nés et résidaient au Burundi au moment de l'enregistrement. Ils sont de formation variée, leur niveau de formation allant des humanités générales au grade de docteur.

Trois des six locuteurs détiennent une licence en linguistique africaine et deux d'entre eux enseignent le Kirundi au secondaire et l'autre est journaliste. Un locuteur est docteur en linguistique africaine et enseignait à l'Université du Burundi au moment de l'enregistrement. Les deux autres locuteurs ont le diplôme de Licence en langue et littérature anglaises ; l'un des deux enseignait l'anglais à l'école secondaire alors que l'autre faisait partie du personnel administratif de l'Université du Burundi au moment de l'enquête. Le choix de ces locuteurs n'a cependant pas été guidé par leur profil, mais par le souci d'avoir des locuteurs qui savent bien lire¹.

Ces données ont été recueillies par l'auteur de ce travail dans le cadre de son travail de fin d'études de Master (Nimbona 2010). Leur transcription s'est réalisée sous le logiciel Praat (Boersma et Weenink 2008) à la lumière des conventions de transcription régissant les corpus de la banque de données VALIBEL².

Les outils informatiques développés pour l'annotation automatique du corpus comme le script EasyAlign (Goldman 2011) du logiciel Praat n'étant pas adaptés à la langue Kirundi, nous avons eu recours à l'annotation manuelle. Dans un premier temps, nous nous sommes d'abord limité à la phase de segmentation et d'alignement du flux de parole en phrases. La segmentation en mots a été réalisée uniquement pour les besoins d'analyse.

Le repérage et l'étiquetage des séquences de type XY (X étant la séquence d'origine et Y la séquence d'autocorrection sur lesquelles a lieu le phénomène d'autocorrection) ont également été réalisés manuellement et le classement des autocorrections s'est réalisé en fonction de la nature morphologique de l'unité corrigée.

L'analyse quantitative des marques du TdF étudiées a été réalisée grâce à plusieurs outils informatiques. Le script *TextGridTools* du logiciel Praat nous a permis de faire des comptages des autocorrections dans chaque fichier annoté sous Praat. Les données extraites de Praat ont été exportées dans le fichier *Excel* qui nous a aidé à faire des examens statistiques sur nos données.

L'analyse qualitative a consisté à examiner les mécanismes d'autocorrection en fonction de la catégorie morphologique du réparateur, c'est-à-dire la nature morphologique de l'unité corrigée

Résultats

Cette section présente les résultats de l'analyse du sous-corpus de parole lue comparés à ceux de l'analyse du sous-corpus de parole spontanée qui ont été analysées avec la même démarche pour rendre compte de la corrélation entre le taux de production des marques de TdF en général et des autocorrections en particulier avec le degré de préparation du discours.

1. Fréquence des autocorrections selon le style de parole

Le tableau 1 montre la distribution moyenne des marques de TdF selon le style de parole : la lecture et la narration. Les moyennes de distribution sont calculées sur la base des valeurs relatives de chaque marque (y compris la pause silencieuse) pour chaque locuteur rapportées à la totalité des occurrences de chaque marque dans notre corpus.

¹ Le programme de l'école fondamentale fixe le niveau des élèves à la fin de 9ème année au niveau B2 en Kirundi langue maternelle (voir l'introduction générale du Guide de l'Enseignant pour la classe de 9ème année). Cela dit, les locuteurs natifs du Kirundi auront normalement acquis les compétences nécessaires en lecture dans cette langue à la fin de l'école post-fondamentale ou des Humanités générales et de l'Université.

² Acronyme de VARIétés Linguistiques du français en BELgique

Tableau 1 : répartition des marques de TdF dans le corpus selon le style de parole (adapté de Nimbona 2010).

Marque	LECTURE			NARRATION		
	Fréquence relative	moyenne du style (%)	Ecart-type	Fréquence relative	moyenne du style (%)	Ecart-type
vh	8,02	1,34	4,52	13,78	2,30	1,74
av	20,70	3,45	4,16	65,98	11,00	8,87
r	10,34	1,72	3,89	21,44	3,57	2,16
auto-i	48,88	8,15	7,49	37,58	6,26	4,33
auto-d	4,67	0,78	3,03	3,58	0,61	2,42
Total	92,61			142,36		

Il apparaît clairement que les marques de TdF en général sont plus fréquentes en narration qu'en lecture avec un total de 142,36 en termes de valeurs relatives, soit près du double du total observé en lecture (92, 61). Nous remarquons cependant un comportement différent des autocorrections immédiates en particulier par rapport à d'autres marques de TdF en fonction du style de parole.

Les autres marques de TdF ont une fréquence faible en lecture et une fréquence forte en narration, alors que, contrairement à celles-ci, les autocorrections sont plus fréquentes en lecture qu'en narration avec une moyenne de 8,15% en lecture contre 6,26% en narration. En première analyse, on peut dire, sur base de ces données, que les autocorrections – en Kirundi – se vérifient moins à l'hypothèse selon laquelle moins un discours est préparé, plus il y a des hésitations.

2. Classe morphologique des autocorrections

Le tableau 2, ci-dessous, indique la répartition des autocorrections immédiates dans la parole lue et dans la parole spontanée (narration) en fonction de la catégorie morphologique que la marque affecte (mots outils (MO), mots pleins (MP)). Les occurrences sont présentées en termes de valeurs absolues.

Tableau 2 : répartition des autocorrections immédiates dans le corpus et par style en fonction de la classe de l'unité lexicale

Unité lexicale ³	Lecture	Parole spontanée	
	Nombre d'occurrences	Nombre d'occurrences	Total (ensemble du corpus)
MO	23	32	55
MP	41	26	67
total	64	58	122

Il apparaît que les autocorrections (immédiates) ne se comportent pas de la même façon en lecture et en parole spontanée. En lecture, les autocorrections immédiates affectent beaucoup plus la catégorie des MP que celle des MO (41 occurrences contre 23), alors qu'en parole non préparée, ce sont les MO qui semblent le plus concernés par les autocorrections que les MP (32 occurrences contre 26). Aussi, nous constatons que, même dans la catégorie des MP, ce sont les verbes qui sont le plus touchés par les autocorrections immédiates.

3. Différents mécanismes d'autocorrection

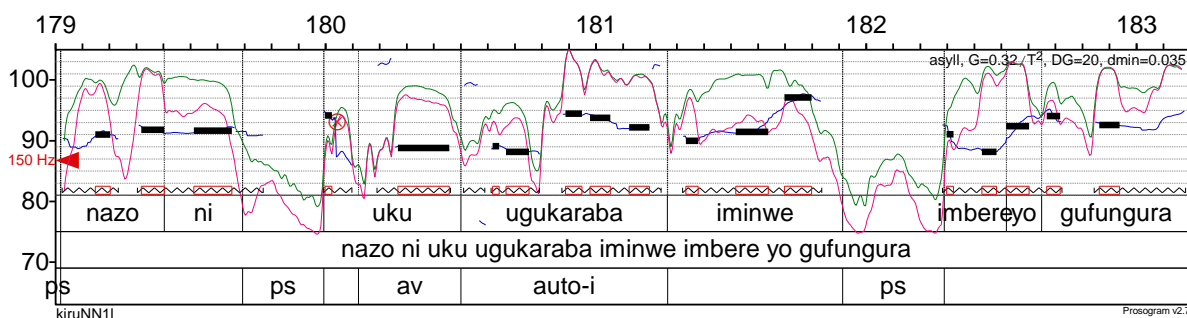
Nous avons remarqué que, dans notre corpus, les autocorrections immédiates peuvent porter sur les traits phonétiques, lexicaux, morphologiques et morphosyntaxiques. Cependant, les mécanismes d'autocorrection ne sont pas les mêmes pour les MO et pour les MP. Pour les MO, nous constatons que les processus de réparation semblent systématiques comme on l'a attesté pour les corpus de français (Candea 2000; Dister 2007). La réparation sur les MO peut

³ MO=mot outil, MP=mot plein

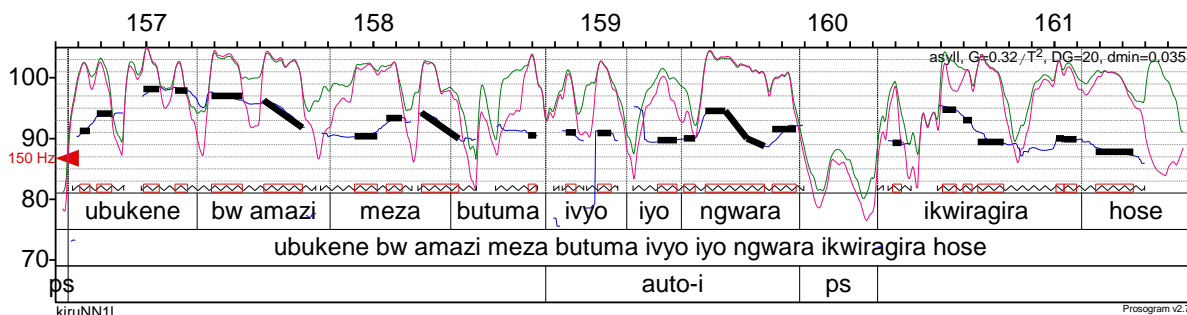
porter sur le trait phonétique ou sur le trait morphologique du nombre, de la personne et du temps tandis que, dans la majorité des autocorrections réalisées sur les MP, la réparation porte sur le trait lexical ou morphosyntaxique.

3.1. Processus de réparation sur les MO

Les autocorrections portant sur les MO où la correction concerne le trait phonétique relèvent du processus d'anticipation phonétique, mais le processus de réparation paraît suivre une loi phonétique dite de *loi de Dahl* particulièrement observée en la linguistique africaine. D'après cette règle phonétique, lorsque deux consonnes sourdes se suivent dans un même mot, la première change et se modifie en une consonne sonore correspondante comme l'on peut le remarquer pour le cas de /k/ et de /g/ sur la bande de prosogramme reproduit en exemple ci-dessous. L'autocorrection a lieu sur le MO, précisément sur le préfixe – ku – de classe 14. Le lecteur, devant le mot lexical ugukaraba « se laver les mains », anticipe sur le mot et commence le mot par le préfixe – ku – normalement caractéristique des verbes à l'infinitif. Mais, par effet de la règle de la langue ci-avant énoncée, le lecteur finit par s'autocorriger et remplacer – ku – par – gu –. L'allongement du premier élément suivi par l'autocorrection du second terme traduit aux yeux de l'auditeur un signe du travail de formulation du locuteur.



Il est important de remarquer que l'autocorrection des MO portant sur le trait morphologique n'entraîne pas le changement de classe grammaticale. La bande du prosogramme suivant montre un cas dans notre corpus où l'autocorrection sur des pronoms démonstratifs **ivyo** « ces » et **iyoy** « cette » qui, dans leur structure morphologique, comportent des préfixes de classe relevant des classes nominales différentes. Le premier ivyo/i-bi-o/ « ces » comporte le préfixe – bi – de classe 8 (s'applique généralement aux choses comptables) tandis que le second iyo/i-i-o/ « cette » comporte le préfixe – i – de la classe 5. Le lecteur lit ivyo « ces » à la place de iyo « cette », peut-être par anticipation phonétique, mais remarque que ce premier pronom est inadéquat devant le mot (i)ngwara « (une) maladie » et s'autocorrige immédiatement. Les deux pronoms étant tous de la catégorie des démonstratifs.



3.2. Processus de réparation sur les MP

Dans l'examen des processus de réparation des MP, nous avons deux types d'autocorrection : les autocorrections lexicales et les autocorrections morphosyntaxiques. Le premier type d'autocorrection s'observe sur deux dimensions différentes : une partie des autocorrections lexicales relève d'un processus d'anticipation phonétique tandis qu'une autre relève du jeu de la tonalité. Dans les deux cas, on observe la substitution d'un MP par un autre. Le second type d'autocorrections, quant à lui, relève uniquement du jeu de tonalité et de la durée.

3.2.1. Les autocorrections lexicales

Le premier type d'autocorrections lexicales peut être interprété de deux façons. Dans l'exemple suivant, le lecteur lit **amabombe** « les bombes » à la place de **amabombo** « les robinets ». D'une part, on peut dire qu'il s'agit d'une anticipation par effet de voisinage phonologique qui existe entre ces deux termes. D'autre part, il y a lieu de penser qu'il s'agirait d'une représentation psychologique ou sociologique du locuteur : le lecteur, par anticipation, prononce non pas ce qui est écrit mais ce qui est en rapport avec la notion de guerre « **ingwano** » qu'il a lue précédemment dans l'énoncé ou avec ce qu'il se représente de cette notion, étant donné l'époque longue et douloureuse que le pays a passé dans la guerre civile.

Ex : kirME21 (188,9 à 197,5) : ... /(671) **ingwano** (la guerre) yahitiye (qui a eu lieu) mu (dans) gihugu (pays) cacu (notre) / (353) yatumye (a fait que) amabombe (les bombes) amabombo (les robinets) rusangi (communautaires) yibwa (soient volés).

[Traduction : La guerre qu'a connue notre pays / (353) a fait que **les bombes les robinets** communautaires soient volés].

Le second type d'autocorrections lexicales que nous observons est quant à lui marqué par la tonalité. Le lecteur produit un item lexical avec un complexe tonal qui n'est pas requis par le contexte syntactico-sémantique. Tout se passe comme si pour le lecteur d'un texte Kirundi, la production va de pair avec l'encodage. Contrairement au cas précédent, bien que nous n'ayons pas des statistiques, il nous semble que ce type d'autocorrection lexicale touche les verbes plutôt que les noms ou les adjectifs. Dans l'exemple qui suit, nous observons la substitution d'un MP marquée par le changement du ton lexical, le segmental étant par ailleurs le même aussi bien pour la séquence d'origine que pour la séquence corrigée si bien que le lecteur étranger à la langue croirait à la répétition.

Ex. kiruNN11 (261 à 268) : abakuru b'a b'amakaritiye nabo //(332) babasaba **gutuza** //(985)
(/les chefs/ de de/ les quartiers/quant à eux //(332) on leur demande de dévaliser //(985) **gutuza** (de chasser/combattre) / abo ceux-là) bose (tous) // (443) ba ba (qui qui) badandariza (qui vendent) imfungurwa (les aliments) ku sur mabarabara (les routes) //

Traduction : quant aux chefs de quartiers //(332), on leur demande de **dévaliser** //(985) **de chasser/combattre** ceux qui font le commerce des aliments sur les routes.

La séquence d'origine qui est la séquence à corriger se prononce normalement avec un ton descendant [HB] sur la voyelle du radical [gutúuza] tandis que la séquence d'autocorrection, celle correcte dans le cas précis, se prononce avec un ton haut sur la voyelle du radical [gutúza].

Nous remarquons en outre que la pause silencieuse (985 ms) qui sépare les deux séquences correspond au temps que le locuteur a mis pour s'apercevoir de l'erreur et de la corriger de par le contexte à la fois syntaxique et sémantique du mot. Cette recherche lexicale aboutit à la reformulation du mot déjà énoncé sous une tonalité autre, ce qui se répercute sur le sens du mot et de la phrase entière.

3.2.2. Les autocorrections morphosyntaxiques

Un autre cas qui nous paraît spécifique à la langue Kirundi (et peut-être aussi à d'autres langues de la même typologie) est le cas d'autocorrections qui portent sur le trait morphosyntaxique où nous observons un passage d'un mode à un autre marqué par l'effet de la tonalité comme pour le cas que nous venons d'évoquer dans le paragraphe précédent. Dans l'exemple qui suit, nous assistons à un passage du mode relatif au mode participe.

Ex2. kiruNN11 () : ... aranze/⁴ euh aranagenzura hose //(503) **yigisha yigisha**
(/ ? / euh il inspecte / un peu partout // (503) **qu'il enseigne en enseignant**
/ abanyagihugu (la population) ingene (comment)/bokwirinda (ils se protégeraient) euh bo/ (euh ils/)
bokwikingira (ils se préviendraient) //(990) iyo (cette) ngwara (maladie).
→ Traduction : ...il.../ euh il passe un peu partout //(503) qu'il enseigne *en enseignant* la population
comment elle peut se prévenir //(990) de cette maladie.

La séquence d'origine et la séquence d'autocorrection ont toutes une voyelle longue du radical et une voyelle post-radical longue (du suffixe – īsh –). La séquence d'origine se prononce toutefois avec une tonalité basse sur la voyelle du radical et avec une tonalité descendante sur la voyelle post-radical [jìgîṣà]. La séquence corrigée par contre et qui est requise par le contexte se prononce avec une tonalité descendante sur la voyelle du radical et une tonalité basse sur la voyelle post-radical [jîgîṣà]. Il y a eu un inter-changement tonal et le lecteur finit par reprendre l'énoncé pour la redresser avec le mot qui s'avère adéquat au contexte.

Tout se passe comme si pour le lecteur d'un texte Kirundi, la production va de pair avec l'encodage. Il n'est pas aisé dans sa lecture comme c'est le cas pour le lecteur du français où les mots sont toujours notés avec leurs signes diacritiques. En Kirundi, la tonalité est produite mais elle n'est pas écrite. Le lecteur doit toujours faire appel à la mémoire pour lier le segmental et le suprasegmental de par le contexte à la fois syntaxique et sémantique du mot dans la phrase.

Discussion des résultats

Notre travail s'est penché sur le phénomène d'autocorrection en Kirundi particulièrement dans la parole lue. Notre hypothèse de départ était que les pauses sonores n'apparaissent pratiquement pas à l'oral de lecture. Certaines marques de TdF vérifient cette hypothèse. En effet, à l'exception des autocorrections immédiates, les autres marques de TdF sont faiblement représentées en lecture et connaissent une forte représentation en narration, parole spontanée par excellence. Nous observons une moyenne très élevée des autocorrections à l'oral de lecture, ce qui déroge l'oral de lecture en Kirundi à cette hypothèse largement vérifiée pour les autres langues comme le français (voir Candea 2000).

Nous avons montré que le lecteur d'un texte Kirundi a recours à plusieurs mécanismes de réparation dont certains sont initialement reconnus dans d'autres études pour les corpus oraux (voir Candea 2000 pour le français) et d'autres qui sont spécifiques à la langue Kirundi et aux autres langues de la même typologie que celle-ci. Le jeu de la tonalité qui n'a d'équivalent dans les langues européennes joue beaucoup dans le travail de décodage du texte Kirundi. La méthode d'analyse utilisée est prometteuse ; l'analyse des autocorrections immédiates à l'oral de lecture permet de découvrir les indices susceptibles d'éclairer le fonctionnement de la langue Kirundi.

L'examen des mécanismes de réparation nous a permis de montrer que la lecture du texte Kirundi, plus qu'en français, recèle les processus de décodage et d'encodage et que les

⁴ La traduction de cette séquence nous est difficile dans la mesure nous ne pouvons pas deviner le mot amorcé parce que sa structure n'a rien de commun avec le mot qui suit le euh marque de TdF.

processus de production et d'encodage se réalisent simultanément. Le lecteur doit toujours faire appel à la mémoire pour lier le segmental et le suprasegmental de par le contexte à la fois syntaxique et sémantique du mot dans la phrase. Cela prouve à suffisance que, malgré leur caractère universel, la réalisation des autocorrections comme d'ailleurs toutes les autres marques de travail de formulation peut varier d'une langue à l'autre, de la famille de langues à l'autre et du style de parole à l'autre.

Conclusion

Ce travail présente des résultats intéressants et s'avère original. Il aborde la linguistique de l'oral, un domaine qui jusqu'ici reste moins exploité pour la langue Kirundi. Il s'agit d'un travail exploratoire et est donc très limité. L'analyse a été faite sur un petit corpus et s'est focalisée uniquement sur les autocorrections immédiates ; ce qui constitue une faiblesse. Il serait intéressant d'élargir l'étude des phénomènes de l'oral à de grands corpus portant sur des styles de parole variés allant de la parole lue à des interactions interpersonnelles et à des débats publics pour découvrir davantage les spécificités de la langue orale en Kirundi par rapport à la langue écrite. Malgré cela, ce travail a montré que, malgré le caractère universel des marques de travail de formulation, leur fréquence et leur mode de réalisation peut varier d'une langue à l'autre et de style de parole à l'autre, ce qui constitue sa force.

Remerciements

Je remercie vivement les locuteurs qui nous ont accordé leur voix et qui ont accepté que nous utilisions les données à des fins scientifiques.

Conflits d'intérêt : aucun à déclarer

Utilisation de l'Intelligence Artificielle : aucune

Références

- Bénard, F. 2005. « Normalisation de corpus oraux : des métadonnées à l'annotation des transcriptions ». Mémoire de Maîtrise. Université Paris 3, Sorbonne Nouvelle.
- Boersma, Paulus Petrus Gerardus, et David Josph Maria Weenink. 2008. « Praat: Doing Phonetics by Computer ». 2008.
- Candea, Maria. 2000. « Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané. Etude sur un corpus de récits en classe de français. » Thèse de doctorat nouveau Régime. Université Paris 3 – Sorbonne Nouvelle.
- Candea, Maria. 2001. « EUH et allongements dits « d'hésitation » : contraintes combinatoires ». 2001 14: 35-46.
- Cenoz, Jasone. 1998. « Pauses and communication strategies in second language speech ». 1998: 11.
- Dister, Anne. 2007. « De la transcription à l'étiquetage morphosyntaxique. Le cas de la banque de données textuelles orales VALIBEL ». Thèse. Université catholique de Louvain.

- Dister, Anne. 2008. « L'autocorrection immédiate en français parlé : le cas des déterminants ». In *Journées internationales d'Analyse statistique des Données Textuelles*.
- Duez, Danielle. 1999. « La fonction symbolique des pauses dans la parole de l'homme politique, in Faits de langues, ». 1999 (13): 91-97.
- Duez, Danielle. 2001. « Caractéristiques acoustiques et phénomènes des pauses remplies dans la conversation en français ». 2001 20: 31-48.
- de Fornel, Michel, et Jean-Marie Marandin. 1996. « L'analyse grammaticale des auto-réparations ». 1996: 8-68.
- Goldman, J.-P. 2011. « EasyAlign: an automatic phonetic alignment tool under Praat ». In *Interspeech'11, 12th Annual Conference of the International Speech Communication Association*, Firenze.
- Goldman-Eisler, Frieda. 1972. « Pauses, Clauses, Sentences ». *Language and Speech* 15(2): 103-13. doi:10.1177/002383097201500201.
- Goulian, Jérôme, et Jean-Yves Antoine. 2001. « Compréhension Automatique de la Parole combinant syntaxe locale et sémantique globale pour une CHM portant sur des tâches relativement complexes ». In *TALN 2001 Recital 2001 (Tours, 2-5 juillet 2001)*., Tours, 203-12.
- Grosjean, F., et A. Deschamps. 1972. « Analyse des variables temporelles du français spontané ». 1972: 129-49.
- Grosjean, François, et Alain Deschamps. 1975. « Analyse Contrastive Des Variables Temporelles de l'anglais et Du Français: Vitesse de Parole et Variables Composantes, Phénomènes d'hésitation ». *Phonetica* 31(3-4): 144-84. doi:10.1159/000259667.
- Henry, S. 2005. « Quelles répétitions à l'oral ? Esquisse d'une typologie ». In *La Linguistique de corpus*, Rennes: Presses universitaires de Rennes, 81-92.
- Henry, S., E. Campione, et Jean Véronis. 2004. « Répétitions et pauses (silencieuses et remplies) en français spontané ». In *Actes des 25es Journées d'Étude sur la Parole (JEP'2004) (19-22 avril, Fès)*, Fès, 261-64.
- Levelt, W. 1983. « Monitoring and self-repair in speech ». *Cognition* 14(1): 41-104. doi:10.1016/0010-0277(83)90026-4.
- Nimbona, Gélase. 2010. « Pauses et hésitations et leur distribution syntagmatiques en Kirundi ». Mémoire de Master. Université catholique de Louvain.
- Passot, F. 2004. « La hiérarchisation des constituants discursifs dans un corpus d'anglais oral spontané ». Thèse de doctorat. Université Paris 3 – Sorbonne Nouvelle.
- Schegeloff, E.A., G. Jefferson, et H. Sacks. 1977. « The Preference for Self-Correction in the Organization of Repair in Conversation ». 1977 53(2): 361-82.

Shriberg, E. 1999. « Phonetic consequences of speech disfluency ». In *Proceedings of the XXIVth International Congress of Phonetic Sciences*, San Francisco, 619-22.

Valli, André, et Jean Véronis. 1999. « tiquetage grammatical des corpus de parole : problèmes et perspectives ». 1999 2(IV): 113-33.