

Le schwa final en français standard est-il un « lubrifiant phonétique » ?

Mathilde Hutin¹, Adèle Jatteau², Ioana Vasilescu¹, Lori Lamel¹ and Martine Adda-Decker^{1,3}

¹Université Paris-Saclay, CNRS, LIMSI, 91400, Orsay, France

²Université de Lille, CNRS, UMR 8163, STL, Lille, France

³Université Paris 3 / Sorbonne Nouvelle, CNRS UMR 7018, LPP, Paris, France

Résumé. Le schwa est une voyelle faible ou réduite notée [ə] alternant avec zéro et restreinte aux syllabes non-accentuées. En français standard, il peut faire surface à l'intérieur ou en fin de mot. Nous proposons ici une étude du schwa final de mot exclusivement, en particulier par le prisme de la question du schwa final en tant que « lubrifiant phonétique » (Purse 2019). Le schwa final est-il réellement un lubrifiant ? Joue-t-il seulement un rôle sur le plan exclusivement phonétique ? Pour répondre à ces questions, nous avons utilisé trois très grands corpus du français (plus de 110 heures de discours) pour établir la présence du schwa final selon les contraintes phonotactiques (la loi des trois consonnes, Grammont 1894) et le style de parole, mais aussi son rôle sur les phénomènes d'adjacence de bas niveau que sont le dévoisement final et l'assimilation régressive de voisement en français standard. Nous concluons que le schwa final est en effet corrélé au nombre de consonnes dans la séquence, et au style de parole ; de surcroît, sa présence est significativement corrélée à beaucoup moins d'effets d'adjacence – comme s'il jouait le rôle de bouclier, facilitant l'adéquation entre forme de surface et forme sous-jacente.

Abstract. Is word-final schwa in Standard French a “phonetic lubricant”? Schwa is a weak or reduced vowel noted [ə], alternating with zero and limited to unstressed syllables. In Standard French, it can appear word-internally or word-finally. In this paper, we present an extended study of word-final schwa, and more precisely about its role as a “phonetic lubricant” (Purse 2019). Is word-final schwa really a lubricant? Does it indeed play a role at a purely phonetic level? To answer these questions, we used three large corpora of French containing more than 110 hours of speech to establish the presence of word-final schwa as a function of the number of consonants in the sequence (Three Consonants Rule by Grammont 1894) and of speech style. We also assess the role of schwa on low-level adjacency effects such as final devoicing and regressive voicing assimilation. We conclude that word-final schwa is correlated with ease of production, with careful speech more than casual speech, and that its presence is highly correlated with less adjacency effects – as if it acts as a shield blocking consonant alterations, thus facilitating the mapping between surface form and underlying form.

1 Introduction

Le schwa, aussi appelé « e muet », « e caduc », « e instable » ou encore « e féminin » (lorsqu'il est en fin de mot), désigne une voyelle faible ou réduite notée [ə] alternant avec zéro et restreinte aux syllabes non-accentuées. En français standard, il peut faire surface à l'intérieur du mot comme dans [səmən] *semaine*, ou [kanətō] *caneton*, ou en fin de mot, comme dans [katʃə], *quatre*. Depuis Mende (1880) et surtout Grammont (1894), les schémas de réalisation ou non-réalisation du schwa ont été largement étudiés (Martinon 1913, Leray 1930, Malécot 1955, 1976, Fouché 1956, Delattre 1966, Dell 1970, Anderson 1982, Côté 2002, Racine & Grosjean 2002, Bürki et al. 2007, Wu et al. 2016, 2017 entre autres). Cependant, peu de travaux se penchent en particulier sur sa distribution en fin de mot, à l'exception notable de Dell (1970, 1976) et Brand & Ernestus (2018), qui restreignent néanmoins leurs analyses à schwa après coda branchante et de Fónagy (1989), Hansen (1997, 2003) et Hansen & Mosegaard-Hansen (2003) qui se penchent sur un type de schwa final particulier, dit « schwa prépausal », sorte de balise sociolinguistique spécifique à la fin de phrase en français parisien (qui ne recouvre donc pas exactement ce qu'on entend par « schwa final »).

Cette préférence pour le schwa interne tient sans doute au fait que les études sur le schwa sont nécessairement limitées par les données. Les plus anciennes se contentent d'étudier, à partir d'intuition d'informateurs (souvent l'auteur lui-même), les endroits où un schwa peut se réaliser ou non. Elles sont par conséquent logiquement restreintes aux schwas lexicaux, pour lesquels le jugement est clair (par exemple l'informateur francophone sait qu'on peut prononcer *pelouse* variablement [pəluz] ou [pluz] mais pas *blouse* *[bəluz]). Or, au cours de cette dernière décennie, l'accès à des données massives a permis de reformuler nombre de questions linguistiques notamment liées aux phénomènes de variation. Il semblerait donc que ce soit un cadre pertinent pour interroger le statut du schwa final car nous disposons désormais de données incluant de nombreuses occurrences, couvrant différents styles de parole et impliquant plusieurs locuteurs et locutrices. Une telle approche, permise par les grands corpus et l'alignement automatique, a fait ses preuves dans l'analyse détaillée du schwa interne (Bürki et al. 2007, Wu et al. 2016, 2017) ou même du schwa final soit après coda branchante seulement (Brand & Ernestus 2018), soit dans des variétés non-standard (Durand & Eychenne 2004, Ranson & Passarello 2012, Eychenne 2019). La seule étude, à notre connaissance, traitant du schwa final de mot après tout type de coda en français standard grâce à des grands corpus est celle menée par Purse (2019). Cependant, cette dernière s'est limitée à deux corpus de parole journalistique (ETAPE et BREF80) et seulement 2 667 occurrences, et n'investigue que deux points : le rapport entre schwa final et <e> graphique et la possible re-syllabification de la coda en attaque du mot suivant. C'est dans cette mouvance que nous interrogeons ici la position finale de mot et le comportement du schwa, mais en observant, entre autres, les contextes de droite plus finement, dans plusieurs styles de parole, et sur des données plus étendues (plus de 110 heures de discours, soit plus de 120 000 occurrences), ce qui offrira un aperçu global appuyé sur la réalité linguistique.

Plus particulièrement, nous souhaitons utiliser cette méthode pour mettre à l'épreuve l'idée reçue que le schwa final en français standard est avant tout un « lubrifiant phonétique » (Purse 2019 à propos de Hansen 1997). Que recouvre ce terme exactement ? Comment se manifeste cette caractéristique ? Le schwa final de mot en français standard peut-il réellement être désigné comme tel ?

Pour répondre à ces questions, nous interrogerons d'abord ce que recouvre précisément une appellation comme « lubrifiant phonétique » (Section 2). Pour répondre aux questions qui en émergeront, nous proposerons une étude statistique du schwa après les douze obstruantes du français /p, t, k, f, s, ʃ, b, d, g, v, z, z/ dans trois grands corpus : ces corpus et la méthodologie suivie seront présentés en détail dans la Section 3. Dans la Section 4, nous

montrons que le schwa final est en effet corrélé à une contrainte phonotactique contre des séquences de trop nombreuses consonnes (4.1) et à un niveau de langage plus soigné (4.2) mais aussi à des phénomènes purement phonétiques (dits « de bas niveau ») comme l'absence des effets d'adjacence que sont notamment le dévoisement final (fortition) et l'assimilation régressive de voisement (lénition) (4.3). La Section 5 sera consacrée à la discussion des résultats et la conclusion de l'article.

2 Qu'est-ce qu'un « lubrifiant phonétique » ?

Cette étude s'intéresse donc à la réalisation, dans l'usage, du schwa en fin de mot. La question que nous nous posons est de savoir dans quelle mesure une qualification en tant que « lubrifiant phonétique » est adéquate pour le désigner. Mais d'abord, qu'est-ce qu'un « lubrifiant phonétique » ?

L'expression est proposée par Purse (2019) au sujet du schwa prépausal étudié par Hansen (1997, 2003) mais ne reçoit nulle part de définition claire. L'expression elle-même suppose deux aspects : un aspect « lubrifiant », et un aspect « phonétique ». Si les deux termes sont ici accolés, il faut se garder de les confondre. Un lubrifiant, peut-on supposer, a pour effet de *faciliter*, d'une façon ou d'une autre, la parole, en production ou en perception. Quant au fait qu'il soit ici qualifié de « phonétique », cela présuppose qu'il joue un rôle sur le plan phonétique sans nécessairement mettre en jeu le niveau phonologique. Notons cependant que des objets ou phénomènes phonétiques ne sont pas nécessairement des « lubrifiants » (par exemple les disfluences, les bégaiements, les métathèses...).

La caractéristique phonétique du schwa final, pour sa part, semble cohérente avec les premières descriptions de ce « ə de détente consonantique », propre au langage emphatique des acteurs ou des orateurs mais qu'on trouve aussi sans mal dans le langage courant (Carton 1999).

Ainsi, Chifflet (1659) observe :

« *e* qu'on appelle féminin (...) est comme le reste d'une consonne qui sonne à la fin d'un mot. Par exemple, *animal* : au bout de cet *l* il y a un petit reste d'un *e*, lequel estant un peu mieux exprimé, l'on entend *animale*. Ainsi *marc*, *marque*, *coc*, *coque*. »

Et Mourgues (1685) d'ajouter :

« les consonnes finales ne peuvent se faire entendre que par le secours d'un petit *e* muet... Quelle différence l'oreille peut-elle percevoir dans la prononciation de *bal* et *bale*, *encor* et *encore*, *vis* et *vice* ? » (cité par Rosset 1911 : 137).

Au 19^e siècle, Rousselot conclut :

« Dans un débit lent et emphatique, il arrive que l'explosion des consonnes sourdes finales devienne sonore et donne naissance à une voyelle » (Rousselot 1897 : 521-523).

Au 20^e siècle, Straka précise :

« En finale absolue (p. ex. dans *mat*, *bac*...), la consonne est articulée avec plus d'énergie qu'en position implosive (p. ex. dans *mat(e)-faim*, *baqu(e)ter*) et parfois même intervocalique..., on entend nettement cette détente, par ex. dans *coq*, *net*, *cep*, *ours*, *mal*, etc. Elle peut donner naissance à une véritable voyelle additionnelle » (Straka 1964 : 24).

Delattre ajoute qu'en français, contrairement à l'anglais :

« la consonne finale se prononce presque comme si elle commençait une nouvelle syllabe : la bouche se rouvre légèrement et un embryon de voyelle se fait entendre. » (Delattre 1966 : 13).

Le schwa final est alors désigné tour à tour comme « le reste d'une consonne », « un petit reste d'un *e* », « un petit *e* muet » ou même « un embryon de voyelle ». Sa caractéristique essentiellement phonétique semble dans ces descriptions faire peu de doute. Mais peut-il être considéré comme un lubrifiant pour autant ?

A cette question s'en superposent plusieurs autres : comment un « lubrifiant » se manifeste-t-il concrètement ? Permet-il de parler plus vite ? Mieux ? Dans ce dernier cas, que veut dire « mieux parler » ?...

Le terme « lubrifiant phonétique » à propos du schwa convoque certainement d'emblée le souvenir de la loi des trois consonnes de Grammont (1894). Cette dernière pose que, dans les cas où l'absence du schwa impliquerait dans la forme de surface une séquence de trois consonnes ou plus d'affilée, alors un schwa est réalisé. Cette loi s'est avérée optimale pour rendre compte de la réalisation du schwa interne (Wu et al. 2016, 2017), mais qu'en est-il du schwa final ? Les « règles » du français canonique supposent effectivement qu'un schwa final devra faire surface pour éviter des séquences de trois consonnes ou plus (Dell 1970), mais qu'en est-il de la réalité langagière ? Dans l'usage, la règle des trois consonnes permet-elle de prédire la réalisation d'un schwa final ? Cette question sera traitée en détail en 4.1 grâce à nos grands corpus.

Ceci étant, le terme « lubrifiant » pourrait aussi renvoyer à l'idée que la parole « lubrifiée » est une parole qui « coule mieux », qui est mieux articulée, plus soignée. Nous n'avons pas dans nos corpus d'étalon de mesure pour observer si une parole est plus soignée qu'une autre. En revanche, le style de langage peut en être un proxy. Nous pouvons considérer la parole journalistique comme plus soignée que la parole informelle, et ainsi observer la quantité de schwas dans nos trois corpus. Ce point sera détaillé en 4.2.

Enfin, le terme « lubrifiant » peut aussi renvoyer à la conception d'une parole pas seulement plus soignée d'un point de vue discursif, mais peut-être aussi d'un point de vue purement phonétique – puisque c'est à ce niveau que le schwa final est censé opérer. Pour interroger le comportement du schwa à cet égard, nous observerons des phénomènes de bas niveau, purement phonétique, et comment le schwa final interagit avec eux. En l'occurrence, bien que le trait laryngal soit distinctif en français, où les non-voisées /p, t, k, f, s, ʃ/ s'opposent régulièrement aux voisées /b, d, g, v, z, ʒ/, il arrive que des « accidents de performance » entraînent le voisement des premières (lénition) ou le dévoisement des secondes (fortition) : il arrive que *départ* soit prononcé [debɑʁ] et *grave*, [gʁɑf]. Or ces réalisations non-canoniques ne sont pas phonologiques : elles arrivent de façon variable et viennent altérer la parole canonique. Nous proposons d'observer les effets du schwa final par rapport à ces productions non-canoniques, dans l'idée que, si le schwa final les empêche (ou en tout cas est corrélé avec moins de productions non-canoniques), nous pourrions proposer une nouvelle définition de ce qu'est un « lubrifiant phonétique ». Dans ce cas, le fait que le schwa final soit corrélé avec moins de voisement accidentel des codas non-voisées et moins de dévoisement accidentel des codas voisées pourrait indiquer que le schwa final de mot permet d'éviter ces altérations phonétiques sur un trait par ailleurs phonologique. Il permettrait alors de réaliser une forme phonétique aussi proche que possible de la forme canonique. Le schwa final serait dans ce cas un lubrifiant phonétique en ce qu'il faciliterait l'adéquation entre forme de surface et forme phonologique. Cette analyse sera présentée en 4.3.

Précisons que nous souhaitons étudier ici si le schwa final peut jouer un rôle de « lubrifiant phonétique » mais qu'il ne s'agit aucunement de l'y réduire. Nous ne comptons pas nous prononcer sur son statut phonologique, ni le comparer systématiquement à schwa interne. La comparaison sera possible pour les points abordés en 4.1 et 4.2, mais elle est impossible pour 4.3. En effet, pour pouvoir observer une quelconque différence entre schwa

interne et schwa final quant aux effets d'adjacence, il aurait fallu observer des schwas internes positionnés de façon à être comparables à un schwa final, c'est-à-dire nous concentrer sur les mots avec un schwa interne soit entre une obstruante voisée et une non-voisée (pour imiter au plus proche la situation où schwa bloque le dévoisement), soit entre une obstruante non-voisée et une voisée (pour imiter au plus proche la situation où schwa bloque l'assimilation régressive de voisement). Malheureusement, malgré la richesse de nos corpus, nous disposons de trop peu d'occurrences remplissant ces critères pour offrir une analyse statistiquement significative.

3 Données et méthode

Pour examiner la présence du schwa final selon le nombre de consonnes environnantes et selon le style de parole ainsi que ses effets sur les phénomènes d'adjacence de bas niveau, nous proposons ici une étude du schwa dans trois grands corpus du français, chacun représentant un style de parole différent (formel, informel et mixte), et de sa réalisation après toutes les obstruantes voisées et non-voisées du français, c'est-à-dire /p, t, k, f, s, ʃ, b, d, g, v, z, ʒ/. Cependant, les [z] de liaison ont été exclus des données. Les sonantes /m, n, ɲ, l, r/ et les approximantes /w, j, ɥ/ ont été laissées de côté car elles ne présentent pas d'alternance de voisement phonologique.

Des phénomènes aussi spécifiques et multifactoriels que le schwa final de mot et la variation accidentelle de voisement peuvent être observés de façon plus fiable grâce à de très grands corpus (Coleman et al. 2016). Outre l'observation à grande échelle qui permet des résultats statistiquement significatifs, l'avantage des grands corpus repose aussi sur le fait que la parole y est plus naturelle que dans des enregistrements opérés lors d'expériences ou d'enquêtes.

Dans cette étude, nous avons utilisé trois grands corpus transcrits manuellement. Le corpus Évaluation des Systèmes de Transcription enrichie d'Émissions Radiophoniques, désormais ESTER (Galliano et al. 2005), comprend originellement 80 heures de discours semi-préparé mais nous l'avons filtré pour enlever les données en français non-métropolitain (RFI et RTM) et ne retenir que les quelque 40 heures en français standard. Le corpus Évaluation en Traitement Automatique de la Parole, désormais ETAPE 1 et 2 (Gravier et al. 2012), contient 13,5 heures de radio et 29 heures de télévision en français, notamment des débats et des conversations. Enfin, le Nijmegen Corpus of Casual French, désormais NCCFr (Torreira et al. 2010), contient 31 heures d'interaction de visu, spontanée, entre amis. Nos trois grands corpus représentent donc chacun un style de parole différent. ESTER est caractéristique du discours préparé, soigné, et potentiellement lu à haute voix ; ETAPE 1 et 2 de monologues semi-préparés et de conversations à deux ou plusieurs interlocuteurs ; et NCCFr de discours spontané entre amis.

Comme ces données ne sauraient être segmentées à la main, elles ont été traitées suivant la méthode décrite dans Gauvain (2002) et Hallé & Adda-Decker (2007, 2011). Un système de reconnaissance automatique de la parole du français a été utilisé pour forcer un alignement pour lequel des variantes avec et sans schwa final étaient possibles. Par exemple, le mot *avec* pouvait être aligné avec les transcriptions [avɛk] ou [avɛkə] selon que la machine estimait qu'un schwa était réalisé ou non à la suite de la vélaire.

Nous avons opéré une vérification manuelle d'un sous-ensemble des données pour vérifier que le schwa final avait été assigné correctement par la machine. La première autrice a écouté 35 occurrences pour chaque coda dans chaque corpus, soit 1260 occurrences, c'est-à-dire à peu près 1% des données. Dès que la transcription était fautive par rapport à la réalisation effective, ou dès que le moindre doute était permis quant à la justesse du jugement, elle a considéré la transcription comme mauvaise. Cette opération nous permet d'estimer que le système de reconnaissance vocale utilisé est fiable à 92,3%.

De plus, des variantes de prononciation ont été introduites pour les codas voisées et non-voisées : le système a été autorisé à sélectionner la coda dite canonique (celle qu'on trouve dans la forme sous-jacente) ou sa variante altérée si la réalisation acoustique correspondait mieux à cette dernière. Par exemple /p/ peut être étiqueté [p] ou [b], et /b/ [b] ou [p]. Ainsi, selon ce que la machine estimait correspondre le mieux à l'audio, un mot comme *soupe* pouvait être aligné avec les transcriptions [sup], [supə], [sub] ou [subə] ; de même, un mot comme *tube* pouvait être aligné avec les transcriptions [tyb], [tybə], [typ] ou [typə].

Grâce à ces trois corpus, nos données comprennent un total de 124 780 obstruantes finales (suivies ou non d'un schwa). Parmi elles, à peu près 70% se terminent avec une coda sous-jacemment non-voisée et 30% avec une coda sous-jacemment voisée. La répartition à travers les trois corpus des codas finales de mots (désormais appelées « codas ») est détaillée dans le Tableau 1.

Tableau 1. Nombre d'occurrences pour chaque coda dans chacun des trois corpus.

	b	d	g	v	z	ʒ	p	t	k	f	s	ʃ	Total
ESTER	832	3901	416	1667	4288	2693	1492	9808	9563	1428	9887	1539	47514
ETAPE	1032	4513	524	2020	5205	2765	1634	9911	9561	1512	9921	1710	50308
NCCFr	509	1839	328	1792	1692	1025	476	7348	6120	814	4346	669	26958
Total	2373	10253	1268	5479	11185	7483	3602	27067	25244	3754	25154	3918	127780

On peut voir aussi, dans la Figure 1 ci-dessous, que la coda la mieux représentée dans tous les corpus est /t/ (21,2%), suivi par /k/ (19,8%) et /s/ (19,7%). La coda la moins représentée est /g/ (1%).

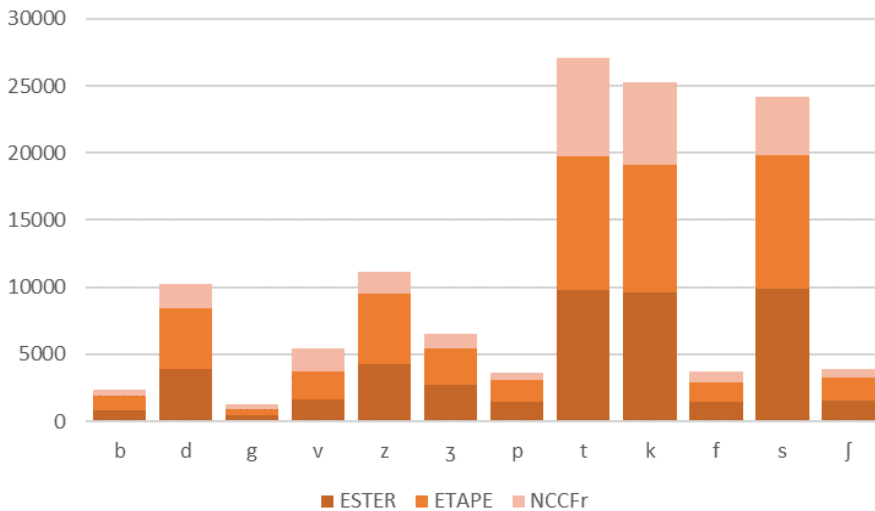


Fig. 1. Nombre d'occurrences de chaque coda dans chaque corpus.

Parmi toutes ces occurrences, on trouve 8071 (17%) schwas finaux dans ESTER, 6283 (13%) dans ETAPE et 2074 (8%) dans NCCFr. Dans son étude, Purse (2019) trouve dans ETAPE et BREF80 un taux de schwas finaux similaire au nôtre (15%).

Ces données nous permettront donc d'observer la présence d'un schwa à travers trois styles de discours, et leurs quelque 110 heures de parole d'observer en détail la fréquence de la réalisation du schwa, et la relation entre la réalisation du schwa final et l'alternance de voisement des codas.

4 Résultats

Observons à présent les résultats de notre investigation pour voir si le schwa final en français est corrélé à la contrainte phonotactique contre les séquences de trois consonnes (4.1) et à un style de langage plus soigné (4.2), puis si les effets du schwa par rapport au dévoisement final (4.3.1) et par rapport à l'assimilation régressive de voisement (4.3.2) sont concluants.

4.1 Le schwa final et la loi des trois consonnes

D'après Grammont (1894), la réalisation d'un schwa devient obligatoire dans la plupart des cas si la forme de surface résultant de la non-réalisation du schwa présente trois consonnes ou plus à la suite (et que ces consonnes ne peuvent pas être re-syllabifiées en une séquence licite). Wu et al. (2017) ont montré que la loi des trois consonnes entre effectivement en jeu dans la réalisation du schwa interne dans les grands corpus. Qu'en est-il du schwa final ?

La Figure 2 ci-dessous montre le taux de réalisations de schwas finaux en fonction de la longueur de la séquence consonantique environnante. Dans nos données, la réalisation du schwa dépend du nombre de consonnes que le locuteur prononcerait à la suite s'il ne réalisait pas de schwa.

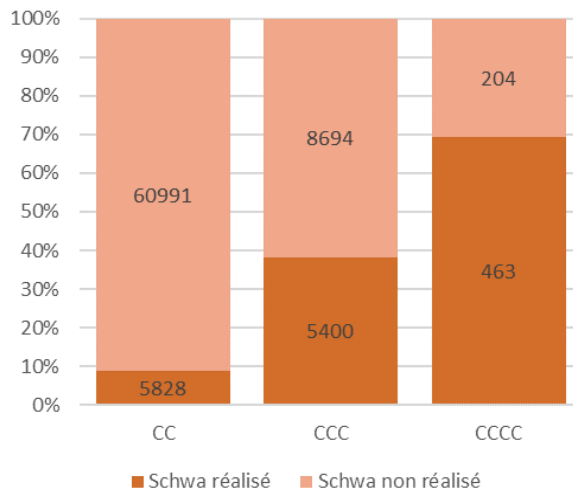


Fig. 2. Réalisation du schwa final en fonction du nombre de consonnes dans les trois corpus

Si la séquence sans schwa fait surface comme une coda simple suivie d'une attaque simple, c'est-à-dire avec deux consonnes de suite (CC), un schwa final est réalisé dans seulement 8,7% des cas. Si la séquence sans schwa fait surface avec trois consonnes de suite (CCC), un schwa final est réalisé dans 38,3% des cas. Notons que dans ce cas, la séquence CCC peut résulter soit d'une coda complexe suivie d'une attaque simple (ex. *Alpes Maritimes*, /alp.maʁitim/) , soit d'une coda simple suivie d'une attaque complexe (ex. *groupe spécialisé*, /gʁup.spesialize/) et que nous n'avons, pour cette étude préliminaire, pas tenu compte de la sonorité (et donc de la possible re-syllabification) de la séquence. Enfin, si la séquence sans schwa fait surface avec quatre consonnes de suite (CCCC), un schwa final est réalisé dans 69,4% des cas. La différence entre les trois groupes est hautement significative ($\chi^2=9963.9$, $df=2$, $p\text{-value}<2.2e-16$).

Il est aussi intéressant de noter que, lorsqu'un item est suivi d'une pause, les codas simples sont moins suivies d'un schwa (25,8%) que les codas complexes (35,9%, $\chi^2=74.802$, $df=1$, $p\text{-value}<2.2e-16$). En cela, le schwa final de mot en français standard se comporte comme le « schwa prépausal » de Hansen & Mosegaard-Hansen (2003).

En conclusion, le schwa final, comme sa contrepartie interne, permet d'éviter des séquences difficiles à prononcer pour un francophone natif : plus il y a de consonnes à la suite dans le cas où schwa n'est pas réalisé, plus un schwa a de chance de l'être. Ces résultats parlent en faveur d'un rôle « lubrifiant » du schwa, puisqu'il fait surface dans les cas où la prononciation serait autrement trop difficile, voire impossible, à l'aune de la grammaire du français.

4.2 Le schwa final selon le style de parole

Dans cette section, nous nous intéressons à la fréquence du schwa, et en particulier si elle est corrélée au style de parole (globalement, formel préparé, formel non-préparé et informel non-préparé), qui est ici un proxy pour une articulation soignée.

Les études passées ont déjà bien démontré l'importante corrélation entre niveau de langue et réalisation du schwa : Wu et al. (2016, 2017) ont montré qu'il y a plus de schwas internes en parole formelle, par exemple le discours journalistique (ESTER), que dans la parole spontanée informelle (NCCFr). Comment le schwa final se comporte-t-il à cet égard ?

Comme nous pouvons le voir dans le Tableau 2, ESTER (discours soigné semi-préparé ou préparé) contient le plus de schwas finaux (17%) tandis que NCCFr (discours informel) en contient le moins (7,7%).

Tableau 2. Nombre de schwas finaux dans chaque corpus.

	Pas de schwa	Schwa	% Schwa
ESTER	39443	8071	16,99
ETAPE	43745	6283	12,56
NCCFr	24869	2074	7,70

Le comportement du schwa final n'est donc, à cet égard, pas différent de celui du schwa interne. La différence entre nos trois corpus étant statistiquement significative ($\chi^2=1292.6$, $df=2$, $p < 2.2e-16$), nous pouvons dire que la présence d'un schwa final est bien corrélée à un niveau de langue plus formel. Le niveau de langue formel étant ici un proxy pour un style de parole mieux articulé, nous pouvons faire l'hypothèse que le schwa final relève bien du « lubrifiant » au sens où il est corrélé à des productions plus soignées.

4.3 Le schwa comme bouclier contre les effets d'adjacence

Observons à présent la question des phénomènes de bas niveau, purement phonétiques, et comment le schwa final interagit avec eux. En l'occurrence, bien que le trait laryngal soit distinctif en français, il arrive que des consonnes voisées soient prononcées dévoisées (phénomène de fortition) et que des consonnes non-voisées soient prononcées voisées (phénomène de lénition). Comme ces réalisations non-canoniques ne sont pas phonologiques, nous proposons que, si le schwa final les empêche (ou en tout cas est corrélé avec moins de productions non-canoniques), nous pourrions proposer une nouvelle définition de « lubrifiant phonétique » comme « tout objet phonétique facilitant l'adéquation entre forme de surface et forme phonologique ».

4.3.1 Schwa vs dévoisement final

À partir d'études sur des grands corpus, Jatteau et al. (2019a,b) montrent que les obstruantes voisées /b, d, g, v, z, ʒ/ peuvent être dévoisées lorsqu'elles sont en position finale de mot devant pause. L'article note que ce dévoisement n'a pas lieu lorsque la coda

est suivie d'un schwa mais n'entre pas davantage dans le détail à ce sujet. Nous nous proposons de nous pencher sur la question puisque, selon notre hypothèse, si un schwa est un lubrifiant *phonétique* au sens où il protège les obstruantes des altérations de bas niveau, ici du dévoisement avant la frontière de mot, alors nous nous attendons à ce qu'il bloque ce dévoisement final.

Pour tester cette hypothèse, nous nous concentrons ici sur les 37 041 occurrences se terminant par une consonne voisée dans la forme de citation. Rappelons que le système d'alignement forcé était autorisé à sélectionner soit l'étiquette « canonique » soit sa contrepartie dévoisée s'il estimait que la réalisation acoustique était plus proche de la seconde que de la première (par exemple, /b/ a pu être étiqueté soit [b] soit [p]).

Les résultats, exposés dans le Tableau 3, montrent que les codas voisées sont significativement moins dévoisées lorsqu'elles sont suivies d'un schwa que lorsqu'elles ne le sont pas ($\Delta=22,7\%$; $\chi^2=1767,8$, $df=1$, $p < 2.2e-16$), avec seulement 12 [Cə]s dévoisés contre 7561 [C]s dévoisés. Ces résultats sont particulièrement nets (encadrés en rouge) lorsque le mot est suivi d'une pause ($\Delta=30,4\%$) – c'est le cas où le schwa bloque le dévoisement final tel que décrit par Jatteau et al. (2019a,b) – mais aussi lorsque le mot est suivi d'une obstruante non-voisée ($\Delta=60,5\%$) – c'est-à-dire qu'il bloque aussi l'assimilation de non-voisement.

Tableau 3. Proportions de réalisations dévoisées de codas voisées en fonction de la présence du schwa final et du contexte de droite.

Contexte de droite	Pause		Voyelle		Sonante		Obstruante voisée		Obstruante Non-voisée	
	_#	_ə#	_#	_ə#	_#	_ə#	_#	_ə#	_#	_ə#
Fin de mot										
ESTER	23,6	0,5	3,8	0,0	4,8	0,0	2,8	0,0	57,8	0,2
ETAPE	27,2	0,3	5,3	0,2	6,3	0,3	5,1	0,0	57,6	0,0
NCCFr	42,3	1,2	11,5	0,0	18,5	0,0	10,0	0,0	66,2	0,0

La Figure 3 ci-après, qui reprend les chiffres exposés dans le Tableau 3, frappe par les différences visuellement impressionnantes entre les barres de gauche, très hautes pour certaines, où la coda est dévoisée alors qu'elle n'est pas suivie d'un schwa, et les barres de droite, toutes extrêmement basses, où la coda est dévoisée alors qu'elle est suivie d'un schwa.

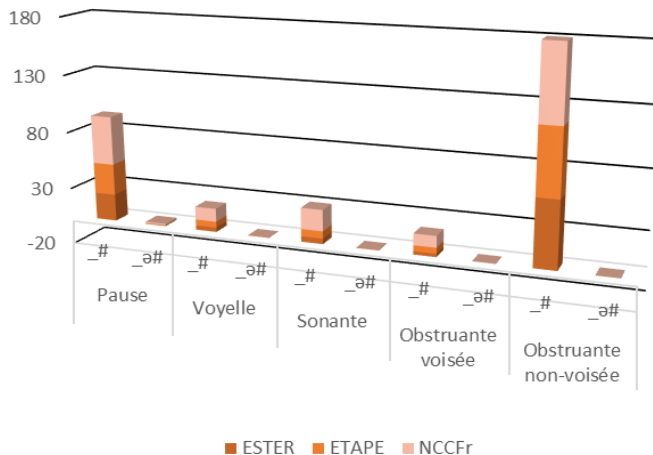


Fig. 3. Nombre de codas voisées subissant un dévoisement final selon qu'elles sont suivies d'un schwa ou non.

Cela signifie que le mot *tube*, par exemple, surtout devant pause ou devant un mot commençant lui-même par une obstruante non-voisée, a moins de chance de subir du dévoisement s'il est prononcé [tybə] que s'il est prononcé [tyb]. Ces résultats sont en accord avec notre hypothèse que le schwa protège les codas du français standard d'un dévoisement final par ailleurs nullement phonologique.

4.3.2 Schwa vs assimilation de voisement

Pour renforcer ces premiers résultats, observons à présent un autre effet d'adjacence : l'assimilation de voisement. En français standard, /p, t, k, f, s, ʃ/ en coda peuvent subir une assimilation régressive de voisement lorsqu'ils sont suivis d'une obstruante voisée (Hallé & Adda-Decker 2007, 2011). Par exemple, la séquence *Philippe Bouvet* peut être prononcée [filibbuve]. Nous faisons l'hypothèse que le schwa protègerait aussi ces codas de cette assimilation.

Pour tester cette hypothèse, nous ne retenons ici que les 87 739 occurrences se terminant par une consonne non-voisée dans la forme de citation. Rappelons que le système d'alignement forcé était autorisé à sélectionner soit l'étiquette « canonique » soit sa contrepartie voisée s'il estimait que la réalisation acoustique était plus proche de la seconde que de la première (par exemple, /p/ a pu être étiqueté soit [p] soit [b]).

Le Tableau 4 montre que les obstruantes non-voisées finales de mot ont tendance à être significativement moins altérées lorsqu'elles sont suivies d'un schwa que lorsqu'elles ne le sont pas ($\Delta=22,3\%$; $\chi^2=2940,8$, $p < 2.2e-16$), en particulier dans le contexte qui provoque l'assimilation de voisement, c'est-à-dire devant obstruante voisée ($\Delta=72,6\%$, encadré en rouge).

Tableau 4. Proportions de réalisations voisées de codas non-voisées en fonction de la présence du schwa final et du contexte de droite.

Contexte de droite	Pause		Voyelle		Sonante		Obstruante voisée		Obstruante Non-voisée	
	_#	_ə#	_#	_ə#	_#	_ə#	_#	_ə#	_#	_ə#
ESTER	3,8	1,2	5,7	0,5	7,1	0,1	70,1	0,2	11,4	0,0
ETAPE	5,3	0,2	11,0	0,9	16,3	0,7	76,7	0,7	19,5	0,4
NCCFr	5,9	0,0	7,2	0,7	11,1	0,0	70,3	0,4	15,4	0,0

Un point intéressant à noter ici est que schwa semble se comporter différemment des autres voyelles (encadrées en pointillés), puisque les voyelles lexicales provoquent un peu de voisement là où schwa semble plutôt le bloquer ($\Delta=6,5\%$, $\chi^2=16.726$, $df=1$, $p=4.318e-05$).

La Figure 4 ci-dessous reprend les chiffres exposés dans le Tableau 4. Elle met elle aussi en évidence les différences entre les barres de gauche, très hautes pour certaines, où la coda est voisée alors qu'elle n'est pas suivie d'un schwa, et les barres de droite, toutes extrêmement basses, où la coda est voisée alors qu'elle est suivie d'un schwa.

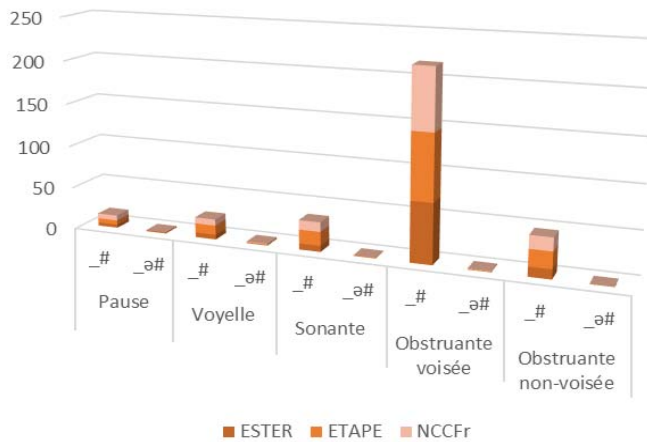


Fig. 4. Nombre de codas non-voisées subissant une assimilation de voisement selon qu'elles sont suivies d'un schwa ou non.

Cela signifie que le mot *soupe*, par exemple, surtout devant un mot commençant par une obstruante voisée, a moins de chance de subir du voisement final s'il est prononcé [supə] que s'il est prononcé [sup]. Une fois encore, ces résultats sont cohérents avec notre hypothèse que le schwa agit comme un bouclier phonétique contre les altérations majeures de la parole.

5 Conclusion et discussion

Nous avons montré dans cette étude que la quantité de schwas finaux est effectivement corrélée à la contrainte phonotactique contre les séquences de trois consonnes. En cela, elle est un lubrifiant qui permet d'éviter des séquences trop difficiles à prononcer ou re-syllabifier selon la grammaire du français. De plus, la quantité de schwas finaux est corrélée à un style de parole plus formel, qui dans notre cas était un proxy pour un style de parole plus soigné. Ce résultat nous invite à considérer le schwa comme un lubrifiant au sens où il permet une parole plus fluide et plus claire. Sur ces deux points, le schwa final n'est pas différent du schwa interne.

Nous avons aussi montré que le schwa final était corrélé à un taux drastiquement bas d'altérations de la forme phonétique comme le dévoisement final et l'assimilation de voisement. Ce résultat nous invite à penser que le schwa final est effectivement un « lubrifiant phonétique » au sens où il est corrélé à des formes de surface plus proches des formes sous-jacentes, ce qui facilite le *mapping* de l'une à l'autre.

Le schwa final en français semble donc bien jouer un rôle de lubrifiant dans les trois définitions que nous en avons donné : sa présence est corrélée à une parole plus facile (il casse les séquences consonantiques) et plus soignée (il est plus fréquent dans les styles de parole plus formels et plus préparés) ; il joue également bien un rôle sur le plan phonétique puisqu'il interagit avec des phénomènes de bas niveau (contre le dévoisement final et l'assimilation de voisement).

Cette étude approfondie du rôle de schwa final en français a tiré profit de l'accès à des grands corpus, des techniques d'alignement automatique et de la méthode statistique. Il fait peu de doute que plusieurs des points que nous avons abordés auraient été impossibles à traiter aussi finement sans eux. Cette étude, donc, au-delà de son intérêt propre quant au comportement du schwa final en français, est aussi une illustration de ce que ce genre de techniques peut apporter à l'étude de la variation.

Cette recherche a été en partie financée par le labex DigiCosme (projet ANR-11-LABEX-0045-DIGICOSME) opéré par l'ANR dans le cadre du programme « Investissement d'Avenir » IDEX Paris-Saclay (ANR-11-IDEX-0003-02).

Références bibliographiques

- Adda-Decker, M. & Lamel, L. (2005). Do speech recognizers prefer female speakers? *Interspeech* 2005: 2205-2208.
- Anderson, S. (1982). The analysis of French schwa: or how to get something from nothing. *Language* 58: 534-573.
- Brand, S. & Ernestus, M. (2018). Reduction of word-final obstruent-liquid-schwa clusters in Parisian French. *Corpus Linguistics and Linguistic Theory* (published online ahead of print). doi: <https://doi.org/10.1515/cllt-2017-0067>.
- Bürki, A., Fougeron, C. & Gendrot, C. (2007). On the categorical nature of the process involved in schwa elision in french. *Interspeech*, août 2007, Antwerp, Belgique.
- Carton, F. (1999). L'épithèse vocalique en français contemporain : étude phonétique. *Faits de langues*, n°13, Mars 1999. Oral-Ecrit : Formes et théories. pp. 35-45
- Chifflet, L. (1659). *Essay d'une parfaite grammaire de la langue française*. Anvers : Jacques van Meurs.
- Côté, M.-H. (2000). *Consonant Cluster Phonotactics: A Perceptual Approach*. PhD-thesis, MIT.
- Coleman, J., Renwick, M., & Temple, R. (2016). Probabilistic underspecification in nasal place assimilation. *Phonology*, 33(3), 425-458.
- Delattre, P. (1966). *Studies in French and comparative phonetics*. La Haye : Mouton.
- Dell, F. (1970). *Les règles phonologiques tardives et la phonologie dérivationnelle du français*. PhD thesis, MIT.
- Dell, F. (1976). Schwa précédé d'un groupe obstruante-liquide. *Recherches Linguistiques* 4. Université de Paris 8-Vincennes, 75-111.
- Dell, F. (1982). On gaps undetectable for language learners. In J. Mehler, M. Garrett & E. Walker (eds.). *Perspectives in Mental Representation*. New York : Erlbaum, pp. 431-438.
- Durand, J. (2014). À la recherche du schwa : données, méthodes et théories. *Congrès Mondial de Linguistique Française – CMLF 2014*, 23-43.
- Durand, J. & Eychenne, J. (2004). Le schwa en français : pourquoi des corpus ? *Corpus* n°3 « Usage des corpus en phonologie », 311-356.
- Eychenne, J. (2019). On the deletion of word-final schwa in Southern French. *Phonology*, 36.3, 1-35.
- Fónagy, I. (1989). Le français change de visage ? *Revue Romane* 24, 2 : 225-254.
- Fouché, P. (1956). *Traité de prononciation française*. Paris : Klincksieck.
- Galliano, S., Geoffrois, E., Mostefa, D., Choukri, K., Bonastre, J.-F. & Gravier, J. (2005). ESTER Phase II Evaluation Campaign for the Rich Transcription of French Broadcast Newshase II Evaluation Campaign for the Rich Transcription of French Broadcast News. *Interspeech* 2005, 2453–2456
- Gauvain, J.-L., Lamel, Lori & Adda, G. (2002). The LIMSI broadcast news transcription system. *Speech communication* 37 (1-2), 89–108
- Grammont, M. (1894). Le patois de la Franche-Montagne et en particulier de Damprichard (Franche-Comté). IV : La loi des trois consonnes. *Mémoires de la Société de linguistique de Paris* 8: 53-90

- Gravier, G., Adda, G., Paulson, N., Carré, M., Giraudel, A. & Galibert, O. (2012). The ETAPE corpus for the evaluation of speech-based TV content processing in the French language. *LREC Eighth international conference on Language Resources and Evaluation*.
- Hallé, P. & Adda-Decker, M. (2007). Voicing assimilation in journalistic speech. *16th International Congress of Phonetic Sciences*, 2007, 493–496.
- Hallé, P. & Adda-Decker, M. (2011). Voice assimilation in French obstruents: A gradient or a categorical process? *Tones and features: A festschrift for Nick Clements*, De Gruyter, 149–175.
- Hansen, A. B. (1997). Le nouveau [ə] prépausal dans le français parlé à Paris. *Polyphonie pour Ivàn Fónagy*, Paris : L'Harmattan, 173–198.
- Hansen, A. B. (2003). Le contexte prépausal - un contexte dynamique pour le schwa dans le français parisien. *La prononciation du français dans sa variation (Phonologie du Français Contemporain)* 142–144. Paris : Jouve.
- Hansen, A. B. & Mosegaard-Hansen, M-B. (2002). Le e prépausal et l'interaction. In A. B Hansen, A. B. et Hansen, MB. M. (éds.), *Structures linguistiques et interactionnelles dans le français parlé*. Actes du Colloque international, Université de Copenhague du 22 au 23 juin 2001 (Études romanes 54). Copenhague: Museum Tusulanum Press, 89-109.
- Hansen, A. B. & Østby, K. A. (1997). Variation in the capital city of France: Paris. S. Detey, J. Durand, B. Laks, & C. Lyche (ed.), *Varieties of Spoken French*. pp. 173-198
- Jatteau, A., Vasilescu, I., Lamel, L., Adda-Decker, M & Audibert, N. (2019a). “Gra[f]e!” Word-final devoicing of obstruents in Standard French: An acoustic study based on large corpora. *Interspeech* 2019, Graz (Autriche)
- Jatteau, A., I. Vasilescu, L. Lamel & M. Adda-Decker (2019b). Final devoicing of fricatives in French: Studying variation in large-scale corpora with automatic alignment. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia, 295–299.
- Leray, F. (1930). La loi des trois consonnes. *Revue de philologie française* XLII. pp. 161-184.
- Malécot, A. (1955). The elision of the French mute-e within complex consonantal clusters. *Lingua*, 5, 44-60.
- Malécot, A. (1976). The effect of linguistic and paralinguistic variables on the elision of the French mute-e. *Phonetica*, 33, 93-112.
- Martinet, A. (1969). *Le français sans fard*. Paris : Presses universitaires de France.
- Martinson, P. (1913). *Comment on prononce le français*. Paris : Larousse.
- Mourgues, M. (1685). *Traité de la poésie française*. Paris : Guillaume de Luyne.
- Purse, R. (2019). Variable Word-Final Schwa in French: An OT Analysis. *University of Pennsylvania Working Papers in Linguistics*, Vol. 25, Issue 1, Proceedings of the 42nd Annual Penn Linguistics Conference, Article 22
- Racine, I. et Grosjean, F. (2002). La production du E caduc facultatif est-elle prévisible? Un début de réponse. *French Language Studies*, 12, 307-326
- Rosset, T. (1911). *Les origines de la prononciation moderne étudiées au XVIIème siècle d'après les remarques des grammairiens et les textes en patois de la banlieue parisienne*. Paris : Armand Colin.
- Rousselot, J.-P. (1897-1901). *Principes de phonétique expérimentale*, tome I. Paris-Leipzig : Welter.
- Straka, G. (1964). *L'Évolution phonétique du latin au français sous l'effet de l'énergie et de la faiblesse articulaires*. Paris : C. Klincksieck.
- Torreira F., Adda-Decker M. & Ernestus M. (2010). The Nijmegen Corpus of Casual French. *Speech Communication*, Elsevier: North-Holland, 2010, 52 (3)

Wu, Y., Adda-Decker, M. & Fougeron, C. (2016). Rôle des contextes lexical et post-lexical dans la réalisation du schwa : apports du traitement automatique de grands corpus. *Actes de la conférence conjointe JEP-TALN-RECITAL 2016*, volume 1 : JEP. pp. 633-641

Wu, Y., Adda-Decker, M., Fougeron, C. & Lamel, L. (2017). Schwa Realization in French: Using Automatic Speech Processing to Study Phonological and Socio-Linguistic Factors in Large Corpora. *Proc. Interspeech 2017*, 3782-3786