Espaces Linguistiques

Exploring norms and variation in L1/L2 pronunciation

Entre écarts et normes de prononciation : acquisition, variation, apprentissage

Susan MOORE MAUROUX

University of Limoges, CeReS, UR 14922

Maelle AMAND

University of Limoges, CeReS, UR 14922

maelle.amand@unilim.fr

Introduction

Today, many tools are available to enable researchers to conduct in-depth studies of pronunciation characteristics: acoustic analysis tools like Praat (Boersma & Weenik, 2007) showing precise phonetic detail, statistical analysis tools allowing researchers to map patterns and draw relevant conclusions from their results. In this volume, contributions include a wide range of methodologies to address the question of norms and variation, focusing on both native languages (L1) and foreign language learning (L2). The languages under study are English, both as L1 and L2, French as L1 and L2, Japanese L1, and L1 Qeltu, a variety of Arabic. A qualitative analysis approach is favoured with data collection methods varying from lab experiments to cohorts of students and official corpora (the DECTE corpus), the focus being mainly on speech production but one study also includes speech perception. Several dozen participants were involved in each study representing different age groups, from primary school learners to adults (mainly young), both male and female.

Our question "norms and variation" reflects a gradient rather than any absolute value. As teachers, we are faced with a paradox: the necessity of presenting learners with a coherent pronunciation system, and our awareness of its true complexity. We may even speak of the kaleidoscopic nature of variation in language. But as researchers, we are free to consider language in all its diversity. We shall therefore begin from the perspective of variation. Indeed, phonetic analysis immediately highlights the diversity of realizations for apparently similar utterances and when these occur regularly (as opposed to performance-related differences), we may speak of "variation". As Ogden (2012, p. 49) states, "it is a truism to say that speech–especially normal conversational speech–is highly variable". Variation in speech can be studied from a number of perspectives: phonetic, phonological, social… (Chambers *et al.*, 2013; Thomas, 2011). Moreover, if the term "variation" implies diversity, and perhaps individuality, there is also an implicit comparison with some sort of general "norm", raising the question of the degree of acceptability (a social concept) and even intelligibility of such variation (cf. Nyberg, J., & Strömbergsson 2021 in SLT, Gooskens 2017 on dialect intelligibility, Henderson 2021 on intelligibility in the EFL[1] context).

1. Atypical profiles in L1: questioning norms

Intelligibility is particularly significant in a medical context where speech impediments or defects mean that language may deviate from what is considered standard, possibly to the extent that communication is undermined. In such cases, it appears relevant to establish what the

---

1 English as a Foreign Language.

norm/standard is. Some researchers suggest that individual deviation may also be considered of interest for research. In his article "La Richesse des écarts à la norme"[2], Bullinger (2002) analyses child language development and argues that what lies outside the standard also needs to be taken into account. This is also the standpoint of Cattu Alves, Ode & Strömbergsson in their article entitled, "Dealing with the unknown–addressing challenges in evaluating unintelligible speech" (2020) where they question the way standard speech is currently evaluated: "[w]hen investigating the interaction between speech production and intelligibility, unintelligible speech portions are often of particular interest" (Cattu Alves *et al.* 2020), proposing additional metrics beyond the common Percentage of Consonant Correct (PCC) which allow us to measure the proportion of intelligible speech (consonants and syllables).

Likewise, Bullinger (2002, p. 12) suggests that even if the performances of atypical children do not match the norm, they still do not necessarily perturb measures. He further argues that trying to understand the specificities rather than approaching them via a norm is the best way to set up adapted communication strategies[3].

Sofi Strömbergsson developed this idea in her paper, "Navigating the search for "normal" in children's speech / language disorders" (2023)[4], confronting two perspectives: on the one hand, intelligibility as a linguistic concept, leading to potential misinterpretation; on the other, acceptability as a social one, highlighting the need for efficient communication with others within a social network. Like Bullinger, she suggests that, when evaluating norms, measures should include non-verbal aspects in narratives. She also suggests that different perspectives should be combined so as to embrace diversity because extensive language variation seems to be increasingly the "norm" rather than the exception.

This diversity clearly represents quite a challenge for speech and language pathologists who need to adapt their appreciation of what the norm is (e.g. Klein, 2011; Jauer-Niworowska, 2022). But if norms need to be questioned, the goal for a speech therapist is still to be able to assess the extent to which deviant pronunciation hinders intelligibility and therefore communication, and to help the patient overcome the difficulties in their L1 where possible. As Jauer-Niworowska puts it: "[a]

---

2 "La Richesse des écarts à la norme" was quoted in the title of the PhonLim 2022 symposium (Limoges University, 18-19 March 2022: la "richesse des écarts à la norme" : individus atypiques en prononciation L1 / L2). It was translated into English as "Outliers, a goldmine for research".

3 The original quotation: "Prendre en compte ces particularités, ne pas les référer à la norme, nous fait comprendre la richesse et la subtilité de ces conduites. Faire l'économie de cette décentration nous diminue et mutile les personnes qui les produisent. Tenter d'en connaitre les mécanismes permet de mettre en place des moyens de soin adaptés." (Bullinger, 2002, p. 103)

4 Sofi Strömbergsson was key speaker at the PhonLim 2023 symposium held in Limoges in June 2023.

message which is not fully functional may regain its functionality after the causes of difficulties in its reception and interpretation and/or correct construction are removed" (2022, p. 9). She defines this strategy as a "reinstatement of normative functioning of the communication" (*Ibid.*).

We cannot but draw a parallel between their objectives and those of L2 teachers who aim to promote communication skills in L2: "[e]nseignants et orthophonistes sont en quelque sorte des professionnels de la norme"[5], remarks Py (2000). He adds: "leur métier consiste en partie à évaluer et à corriger et que l'activation de ce rôle social dans des interactions avec des partenaires novices entraîne (sauf décision contraire) l'établissement implicite d'un contrat didactique"[6]. This raises the question of what "norm" to target as a reference model for learners. Angelovska and Hahn remind us that "in the traditional view of foreign language learning and teaching, the native speaker criterion is taken as a measure of success in learning, as well as a role model for language teaching and as a measuring stick in research" (2009, p. 164). But what is meant by a "native speaker criterion"?

2. Norms, variation and reference

If "nativeness" in EFL generally refers to an aimed target associated with standard pronunciation, native speakers of the same language are far from sharing identical phonetic features. When features are shared and accepted by a recognizable (social) group (Eckert, 1989), they may constitute a new variety (Hickey, 2003) which is therefore regarded as a "norm" by that group of speakers, be it covert or overt (Trudgill, 1972). As research in sociolinguistics has shown, beyond regional variation, there are many layers to the language a native speaker uses, influenced by social environment, register and phonostyle (Léon, 2000), among others.

Despite such diversity, the choice of an L1 reference model for L2 tends to be fairly restricted. It may simply depend on the teachers' own norm which is then passed on to the learners who consciously or unconsciously adopt it (Dalton-Puffer *et al.*, 1997). Reference works tend to refer to a limited number of models, often just one (cf. Michari, 2023, for French textbooks for EFL). In this volume where L2 French and L2 English are studied, reference is made to two varieties of the language and not just one (Kamiyama *et al.*, Capliez), thereby including the concept of variation when dealing with second language learning. For L1 French, these are standard *Parisian* French, aka. *Reference French* (Detey *et al.*, 2016), and Canadian French (used here for research purposes

---

5 "Teachers and speech therapists need to deal with norms as professionals" (our translation).
6 "Part of their job involves assessing and correcting and so this social role, when interacting with novice partners (unless otherwise stated), means establishing an implicit didactic contract" (our translation).

rather than as a teaching model). For English, there have been several ways to refer to British English (Received Pronunciation (RP), BBC English, Standard or Southern British English (SBE)– used as a teaching model by Vaissière & Exare in the present volume. The other variety is generally referred to as General American (GA or GenAm).

2.1. Reference models: what's in a name

The names chosen for English are worthy of note. The term General American (GA) is used in British English pronouncing dictionaries which propose a single American variant corresponding to the British English entry (or entries) and designates what aims to be a commonly acceptable pronunciation that does not single out any social or regional variety (Jones, 2006, p. vi). Yet, in America, the term has been subject to scrutiny, particularly as regards what it covers (Van Riper, 1986). An alternative to GA is proposed by Kretzschmar who considers that the term General American "implies there is some exemplary state of American English from which other varieties deviate" (2008, p. 42). He contests this and proposes to use the term Standard American English (StAmE) to refer to an unmarked pronunciation which "designates the level of quality (here of pronunciation) that is employed by educated speakers in formal settings" (Kretzschmar, 2008, p. 37).

For British English, a number of names have been used and there has been much discussion as to how to refer to a "standard" and what the standard represents. Unlike Kretzschmar's interpretation of "standard" for American English, the association of the term to British English may suggest some sort of recommendation, potentially implying that other varieties are not recommendable. By contrast, Southern British English[7] suggests that varieties are categorised based on a geographical scale rather than a hierarchical one. The discussion is reflected in the explanations given in pronunciation dictionaries for the model described in the entries. The lexicographers' objective is to provide learners with a generally acceptable model; they may also include alternatives and even refer to another variety[8], but their focus is on one (or two) reference models. To designate the variety described in the 12th edition of the *English Pronouncing Dictionary*, Jones refers to Received Pronunciation (RP) "for want of a better term" (Jones, 1963, p. xvi) but qualifies his use of the term as follows: "I should like it to be understood, however, that RP means merely 'widely

---

7 This is the terminology used in recent Board of Examiner's reports for the *Agrégation Externe d'anglais*, a French teaching exam, henceforth referred to as AEA.

8 Today, in both the *Longman Pronouncing Dictionary* (Wells, 2008) and the *English Pronouncing Dictionary* (Jones, 2018), systematic reference to both British and American English is made (and a few variants are included, but these cannot be exhaustive). Fifty years earlier, there was no systematic reference to American English in the 12th edition (Jones, 1963).

understood pronunciation' and that I do not hold it up as a standard which everyone is recommended to adopt." (*Ibid.*) As he says, this variety is widely understood, but it is not in fact widely used, and he concludes: "There exist countless other ways of pronouncing English, some of them being used by large communities." (*Ibid.*). Nearly fifty years later, in the reference section of the 18th edition of the *English Pronouncing Dictionary*, Catherine Sangster in her contribution entitled "The BBC, its Pronunciation Unit, and 'BBC English'" again refers to RP and to the term 'BBC English', which she considers inappropriate:

> There is not (and never was) an official BBC pronunciation standard; its broadcasters speak English with a range of accents. In spite of this, the term 'BBC English' is often used as a synonym for Received Pronunciation (Roach, 2004) and for the sort of accent described in dictionaries like this one. (Jones, 2011, p. xxix)

She shows how far broadcasting has evolved[9], embracing a very broad approach to pronunciation which reflects current BBC policy as regards diversity: "[…] many accents can now be heard. This is felt to be important for the BBC because it means that the voices that the BBC's broad audience hear in its broadcasts are as diverse as their own voices." (Jones, 2011, p. xxix).

The debate on standard varieties reflects changing attitudes to variation itself (Baratta, 2023). Yet, in a learning environment, whatever the name, teachers will promote one or more reference models.

2.2. Norms and reference models for L2

As there are significant differences in the vowel systems of General American and Southern British English, Vaissière & Exare[10] choose to use a single model as their reference for L2 English since they are promoting a method for learning pronunciation, thereby exposing young learners to one coherent pronunciation system. Furthermore, they justify their choice of SBE[11] as the standard for pragmatic reasons[12] because they consider, following Michari (2023), that this is the most widely taught variety of English in France today, and secondly, that the two speakers who participated in the videos speak this variety.

---

9 She explains that the variety described in the dictionary did indeed correspond to BBC newsreaders in the past but she feels: "this would be better viewed as a by-product of the restricted social group from which BBC employees was drawn at that time, rather than a matter of deliberate policy." (Jones, 2011, p. xxix)

10 Vaissière & Exare were guest speakers at the PhonLim 2022 symposium (Limoges University, 18-19 March 2022) where they presented CleanAccent ©.

11 For them, this refers to "Standard British English".

12 « La variété choisie pour les leçons en classe est *l'anglais britannique standard* pour deux raisons. Tout d'abord, l'anglais britannique est encore, par rapport à l'anglais américain, la variété la plus enseignée en France à l'heure actuelle (Michari, 2023). De plus, les deux enseignantes filmées pour l'expérimentation maîtrisent l'anglais britannique standard. »

Likewise, Capliez adopts SBE[13] as the norm but since he focuses on the pronunciation (or not) of <h> and the glottal fricative /h/, SBE and General American share the same characteristic–he does nonetheless identify specificities of General American /h/. But he points out that h-dropping does occur in some varieties (accents) of English, but it is not regarded as a model for learners because h-dropping indexes marked social identities. Indeed, sociolinguistic research has confirmed the correlation between h-dropping and social factors (Wells, 1982, p. 254). This is probably still very marked in terms of social acceptability despite a much more inclusive approach to diversity.

Reference models are generally used to establish a norm for learners' productions. This is the perspective for much teaching and research, as we see in the contributions of Vaissière & Exare and Capliez. But in one of their experiments, Kamiyama *et al.* take a different approach and try to establish possible correlations between listeners with different L1 varieties and their perception of L2 learners' speech. In their experiment, they ask subjects who speak different reference varieties of French (Quebec and Parisian French) to assess the performances of Japanese learners of L2 French, thereby measuring the impact of the variety on perception.

2.3. Norms and atypical profiles in L2

Research in foreign language learning has focused mainly on groups of individuals that have similar second language perception and production strategies rather than on individuals. Studies of L2 often propose to measure the influence of factors that may contribute to explain or improve second language perception and production. Learning models are developed with regard to the learning strategies of the majority. Yet this type of analysis also raises the question of outliers–the speakers who do not correspond to the pattern–that are also worthy of interest and it is by considering what deviates from the norm that we can fully appreciate second language production and perception in all its diversity.

Brevik and Hellekjaer (2017) focused their study on L2 readers with atypical profiles. They studied a statistically marginal group of readers who performed significantly better when reading their L2 than their L1, thus challenging the premise that L1 and L2 reading competences are correlated. Their work questions the way learners acquire reading strategies and they conclude that:

> Outliers show that we need to rethink L1 and L2 reading instruction, not only to understand why these students read markedly better in English as L2 than in their L1, but also to develop these students as readers […] taking students' interests into account when choosing reading tasks and materials. (Brevik & Hellekjaer, 2017, p. 89)

---

13 Capliez refers to Southern British English.

Likewise, Cauvin (2017) in her work on learners and prosodic patterns identifies an atypical profile type which she calls the "pianist profile", students who perform both very well and very poorly as compared with a native model. She hypothesizes that these students have very specific learning strategies, particularly influenced by written English, and that diagnosing the profile and providing them with relevant tools should help to improve performance.

In all of these studies, researchers are trying to bridge the gap between L1 and L2 and, so far, we have considered norms and references with regard to an L1 model for L2 learners. If some learners present atypical characteristics, they all have in common the fact that they go into the learning process with their own L1 norm, namely cross-linguistic phenomena. We shall now consider what the implications of this are.

3. L1 norms and L2 challenges

Learners not only have to address the challenges raised by the native speaker model chosen as the benchmark, they also need to deal with their own L1 norm which will necessarily influence how they perform in L2. Flege (1995) considers that, as this is their primary reference, it will act as a filter influencing both perception and production of L2.

3.1. L1 norm / speaking L2

As regards production, the characteristics of a learner's L1 will trigger fairly predictable variation with respect to any native model. In his Speech Learning Model, Flege (1995) posits that on a segmental level, three categories may be identified: phonemes that are shared between L1 and L2; phonemes which are absent from the L1 inventory; phonemes that are similar to those in L1. He argues that the similar-but-different phonemes will be most tricky to produce as they tend to be articulated as for L1, unlike new phonemes where there is no interference.

This is one of the hypotheses explored by Capliez in his contribution to this volume since the glottal fricative /h/, absent from the L1 inventory of the French students he is analyzing, does not correspond at all to their L1 code. Kamiyama *et al.* go a step further by considering not one but two varieties of L1 Japanese in order to see how the L1 norm of the Japanese learners affects their pronunciation of L2 French.

3.2. L1 norm / interpreting L2

If it is true that the spelling-sound codes of a learner's L1 will influence their pronunciation of written L2 until they integrate the L2 code[14], it is surely listening that is most challenging because, so often, what they see is not what they expect to hear.

Indeed, phonetic variation in L1, and by extension in the L1 model for L2 learners, can be divided into two categories: (1) word internal variation and (2) variation that occurs across word boundaries in connected speech. In the first case, phonetic processes such as assimilation result in pronunciations that are far from the way words are written: *bonfire*, when pronounced /ˈbɒmfaɪə/ may be misinterpreted as *bomb fire*, which is far from the actual meaning. Likewise, *sandwich* is frequently pronounced /ˈsæmwɪʤ/ (like *Sam widge*) with elision of the intermediate consonant in the /ndw/ cluster followed by an anticipatory assimilation of the labialisation in /w/ and voicing of the final affricate. For both <nd> and <ch>, the general reading rules for English do not apply. The native pronunciation will be even less recognizable for learners because, as the word has been borrowed in many languages including French, they are used to saying and hearing it according to their own spelling-sound code. In the second case, as Huart (2010) points out, the process frequently affects unstressed grammatical words (but not only) in connected speech. Phonetic variation occurs especially in spontaneous speech, in interactive situations, where tempo is rapid. Vowels tend to lose their identity (pronounced as the central vowel schwa), consonants may be elided, making the connection between the written word and its pronunciation unrecognizable, which presents a real difficulty for some learners, e.g./jɔːʃɔːtətraʊzəz/for *your shorter trousers* or *you're short of trousers*. Such realizations are readily intelligible for a native speaker who has an innate understanding of these processes. It is often, however, challenging for learners who do not possess the tools to decode the signal.

Yet, beyond this type of predictable context-dependent variation, there is increasing awareness of the diversity of varieties and the necessity for learners to hear them. Even if emphasis in teaching exams is still laid on the coherence of the teacher's pronunciation[15] (and by implication that of the learner), in recent directives of the Common European Framework for Languages, trying to imitate a native norm is considered to be counter-productive[16]; learners simply need to approximate to a

---

14 Moore Mauroux (2022) demonstrates how learning the English spelling-sound code can benefit learners.

15 The 2018 AEA report suggests that awareness of the specificities of the two reference accents (SBE & GA) should be beneficial to candidates / teachers in acquiring a coherent variety of English (AEA 2018, p. 82).

16 « La maîtrise phonologique du locuteur natif idéalisé a longtemps été un objectif de l'enseignement des langues, l'accent représentant un indicateur de mauvaise maîtrise de la phonologie. L'insistance mise sur l'accent et son exactitude au détriment de l'intelligibilité a nui au développement de l'enseignement de la prononciation. Les modèles

chosen standard, not reproduce it. Angelovska and Hahn challenged the premise that learners' proficiency should be measured by "the native speaker criterion" (2009, p. 164) and today the trend is clearly shifting away from this towards a focus on communication and intelligibility (Cook, 1999; Bøhn & Hansen, 2017). Current policy in language teaching in France reflects this approach, and, just as importantly, French official instructions also identify the cultural importance of being exposed to different varieties and not just the sole reference standards (Bouvet, 2021).

4. L1 norms and variation

What lies outside standard L1 norms can contribute to a greater understanding of language and communication. This is what Ogden (2012, p. 54) demonstrates in his paper entitled "Making Sense of Outliers" where he studies apparently atypical phonetic cues (at least in statistical terms) in native American English speakers, arguing that these in fact contribute to intensifying emphasis (IE). His findings show that segmental cues combine with prosodic cues to convey emphasis and, just as connected speech phenomena are perfectly decoded by native speakers, he explains that even though rare tokens do not occur frequently in the language, they remain meaningful, and should not be disregarded in language studies.

Likewise, much research in sociolinguistics has been devoted to studying different pronunciations as well as lexical / grammatical specificities in different varieties of languages, that is to say, where groups of speakers share common pronunciation characteristics. Researchers may choose to explore variation from a number of different perspectives: ethnic, regional, geographical, social, gender… The amount of data available for analysis varies from one language to another, which necessarily has an impact on the methodologies used, ranging from statistical analyses of large corpora to lab experiments on a smaller scale, as we see in the contributions to this volume.

5. The volume: L1/L2 corpora and analysis methodologies

Where for English, large corpora have been set up in order to analyze variation, in other languages, there has been much less background work. This is certainly the case for Qeltu, a variety of Arabic, which is the focus of **Yaseen**'s article. He starts by questioning an established "norm" for explanations of vowel lowering in Qeltu–that the process is triggered by proximity to guttural sounds–and sets out to explore the process further by setting up a number of sociolinguistic interviews involving some thirty respondents. His analysis brings to light hitherto unreported

---

idéalisés qui ignorent les accents, ne prennent pas en compte les contextes, les aspects sociolinguistiques ni les besoins des apprenants. » (CECRL, 2018, p. 140)

patterns of vowel lowering, suggesting that there are external factors involved in vowel lowering in Qeltu that go beyond the system–internal ones that have been traditionally identified.

Amand's contribution is based on a well-established corpus of Tyneside English, the DECTE corpus (Corrigan *et al.*, 2012). She works from the original transcriptions of some forty-four adult speakers aged between 15 and 80, both male and female, who are mainly from working-class backgrounds. Her analysis is based on variation in four key vowels of Tyneside English. Using advanced statistical tools, combining Multiple Factorial Analysis and cluster analysis (Husson, 2010), to identify sub-groups of speakers, she investigates to what extent the speakers correspond to the norm for this variety of English and how coherent their variety is with regard to these 4 vowels. Her results reveal significant class and gender disparity.

For **Kamiyama *et al.***, L1 variation is studied in relation to learning L2. Their analysis is based on a series of lab experiments targeting L2 French which were realized with Japanese speakers from two distinct regions of Japan: Tokyo (Kanto) and Osaka (Kansai). They measure the formant values of Japanese vowels, then study the perception and pronunciation of French vowels, especially /u/. The objective of the study is to establish how far the variety of L1 Japanese affects performance in French. Moreover, production was assessed by native French speakers using different varieties (Quebec / Paris), adding an interesting layer of complexity to the question of norms and variation.

Likewise, **Capliez** has set up his own experimental corpus to explore how French L1 learners cope with phonemes in the L2 reference norm that do not exist in their own L1. He focuses on the pronunciation of /h/ and analyzes ninety-four subjects, forty-two men and fifty-two women, aged between 18 and 23, measuring their performance in three different recorded tasks (word reading, sentence reading, and spontaneous speech). From a detailed comparison of the different speakers, he shows how performance is related to the type of speech; his findings also highlight distinct gender differences. This exploration of the relation between L1 and L2 raises questions about teaching methods in general, including how best to introduce new phonemes.

This is very much the concern of **Vaissière & Exare** who have set up an innovative system for early learning of English in French primary schools. Their approach draws on a great many references in the fields of both linguistics and didactics. The method is based on the premise that introducing a new language–so a new norm–will be more successful if it is done very early, before vocabulary acquisition and before commonly observed errors have had a chance to stick. The method focuses mainly on the sub-segmental and segmental features that are potentially challenging to French learners of English (vowels, consonants or aspiration). The experiment, so

far carried out in one school, highlights the need for researchers, primary school teachers and also parents to work together.

Conclusion

The present volume was largely inspired by the papers presented during the PhonLim 2022 symposium where we began investigating variation in language acquisition, language learning and native languages, with a special interest in outliers and atypical pronunciation. The theme was further developed at the PhonLim 2023 symposium, from the perspective of speech therapy. Papers and the ensuing discussions led us to take a slightly different angle for this volume, addressing both norms and variation in language.

The volume includes studies relating to four languages (French, English, Arabic and Japanese); the question of norms and variation is addressed from two perspectives: on the one hand, foreign language (L2) teaching / learning where the target language is the reference norm, and on the other, variation within a native language (L1 varieties). The first two articles address the challenges of speaking L2 English, beginning with Vaissière & Exare's account of their early learning experiment with young L1 French learners. The method involves the English pronunciation system as a whole, concentrating particularly on what is potentially tricky, so where the L2 reference norm is different from a French learner's L1 norm. This is followed by Capliez's article which highlights a specific challenge in English (pronouncing /h/) for young French adults. The third contribution, Kamiyama *et al.*, also deals with learning a foreign language (this time, L2 French), from both production and perception perspectives. They also introduce a different angle on norms / variation, questioning how the Japanese learner's L1 variety affects performance in L2. The last two articles look at variation in L1, first with Yaseen's work on Qeltu with a corpus he has set up to investigate traditional perspectives on vowel lowering, questioning generally accepted views and developing understanding of a little researched variety. Finally, Amand looks at variation within a variety of English in the DECTE corpus of Tyneside English, identifying linguistic sub-groups within the variety that are mostly gender defined, and collectively form a continuum from localised to less-localised speech.

The wide range of contributions to this volume underscores the significance of diversity as a social construct reflected in language. As learners, teachers and researchers we can only accept the need for constant questioning of our values and recognise their relativity: one person's norm is another person's variety.

References

ANGELOVSKA Tanya & HAHN Angela, 2009, "English with a native-like accent: an empirical study on proficiency", *in* Alessandro G. Benati (ed.), *Issues in Second Language Proficiency,* London, Continuum International Publishing.

BARATTA Alex, 2023, "Regional accents are OK for teaching but not too regional: A discussion of accent preference in British teaching", *in* Alice Henderson & Anastazija Kirkova-Naskova (eds.), *Proceedings of the 7ᵗʰ International Conference on English Pronunciation: Issues and Practices*, p. 1-10.

BOERSMA Paul & WEENINK David, 2007, *Praat: doing phonetics by computer* [Computer program]. *Glot International*, vol. 5, n° 9/10, p. 341-345. http://www.praat.org/

BØHN Henrik & HANSEN Thomas, 2017, "Assessing pronunciation in an EFL context: Teachers' orientations towards nativeness and intelligibility", *Language Assessment Quarterly*, vol. 14, n° 1, p. 54-68.

BULLINGER André, 2002, « La richesse des écarts à la norme », *Enfance*, vol. 54, p. 100-103. DOI : https://doi.org/10.3917/enf.541.0100

BREVIK Lisbeth M. & HELLEKJÆR Glenn Ole, 2018, "Outliers: Upper secondary school students who read better in the L2 than in L1", *International Journal of Educational Research*, vol. 89, p. 80-91. DOI: https://doi.org/10.1016/j.ijer.2017.10.001

Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer. Volume complémentaire avec de nouveaux descripteurs, eduscol.education.fr/langues-vivantes/actualites/actualites/article/volume-complementaire-du-cecr-avec-de-nouveaux-descripteurs.html, février 2018, accessed 19 Sept. 2023.

CATTU ALVES Miriam Oster, ODE Carina & STROMBERGSSON Sofia, 2020, "Dealing with the unknown – addressing challenges in evaluating unintelligible speech", *Clin Linguist Phon*, vol. 34(1-2), p. 169-184. DOI: 10.1080/02699206.2019.1622787

CAUVIN Evelyne, 2017, *Élaboration de critères prosodiques pour une évaluation semi-automatique des apprenants francophones de l'anglais*, thèse de doctorat, université de Paris VII.

CHAMBERS J. "Jack" K. & SCHILLING Natalie, 2013, *The Handbook of Language Variation and Change*, Oxford, Wiley Blackwell.

COOK Vivian, 1999, "Going Beyond the Native Speaker in Language Teaching.", *TESOL Quarterly*. Vol. 33(2), p. 185-209.

CORRIGAN Karen P., BUCHSTALLER Isabelle, MEARNS Adam & MOISL Hermann, 2012, *The Diachronic Electronic Corpus of Tyneside English*, Newcastle University. https://research.ncl.ac.uk/decte

DALTON-PUFFER Christiane, KALTENBOECK Gunther & SMIT Ute, 1997, "Learner attitudes and L2 pronunciation in Austria", *World Englishes*, vol. 16(1), p. 115-128.

DETEY Sylvain, LYCHE Chantal, RACINE Isabelle, SCHWAB Sandra & LE GAC David, 2016, "The notion of norm in spoken French: Production and perception", *in* Sylvain Detey, Jacques Durand, Bernard Laks & Chantal Lyche (eds.), *Varieties of Spoken French*, Oxford, Oxford University Press, p. 55-67.

ECKERT Penelope, 1989, *Jocks and burnouts: Social categories and identity in the high school*, Teachers' College Press.

FLEGE James Emil, 1987, "The production of new and similar phones in a foreign language: Evidence for the effect of equivalence classification", *Journal of Phonetics*, vol. 15, p. 47-65.

FLEGE James Emil, 1995, "Second Language Speech Learning: Theory, Findings, and Problems", *in* Winifred Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, Timonium, York Press, p. 233-77.

FLEGE James & BOHN Ocke-Schwen, 2021, "The Revised Speech Learning Model (SLM-r)", *in* Ratree Wayland (ed.), *Second Language Speech Learning, Theoretical and Empirical Progress,* Cambridge, Cambridge University Press, p. 3-83. DOI: 10.1017/9781108886901.002

HENDERSON Alice, 2021, "Intelligibility and identity: From teaching pronunciation to training for spoken language variation. Education", Université Savoie Mont Blanc ⟨tel-03295473⟩.

HUART Ruth, 2010, *Nouvelle Grammaire de l'anglais oral*, Paris, Ophrys.

HICKEY Raymond, 2003, "How do dialects get the features they have? On the process of new dialect formation", *in* Raymond Hickey (ed.), *Motives for language change*, Cambridge, Cambridge University Press, p. 213-243.

HUSSON François, JOSSE Julie, & PAGÈS Jérôme, 2010, "Principal Component Methods Hierarchical Clustering Partitional Clustering: Why Would We Need to Choose for Visualizing Data?", *Technical report of the Applied Mathematics Department* (*Agrocampus*), http://www.sthda.com/english/upload/hcpc_husson_josse.pdf

JAUER-NIWOROWSKA Olga, 2022, "Reflection on the concept of "the norm" in speech-language therapy / Rozważania nad normą w logopedii", *Logopedia Silesiana*, vol. 2(11), p. 1-22.

KLEIN Alexandre, 2011, « Approche philosophique de l'éthique en orthophonie Un plaidoyer pour une discipline orthophonique : Étude généalogique de l'orthophonie et de ses enjeux épistémologiques et éthiques », *Rééducation orthophonique*, vol. 247, p. 7-24.

JONES Daniel, 1963, *English Pronouncing Dictionary*, 12th ed., London, Everyman.

JONES Daniel, 2011, *English Pronouncing Dictionary*, 18th ed., ed. by Peter Roach, Jane Setter & John Esling, Cambridge, Cambridge University Press.

KRETZSCHMAR William A. Jr., 1996, "Standard American English pronunciation", *in* Edgar W. Schneider (ed.), *The Americas and the Caribbean*, Walter de Gruyter, p. 37-51.

LÉON Pierre, [1993, Nathan Fac] 2000, *Précis de phonostylistique*, Paris, Armand Colin.

MICHARI Dylan, 2023, « Oser dire le nouveau monde : l'enseignement-apprentissage des accents en classe d'anglais à l'aune du plan langues vivantes », *in Quels objectifs pour une phonologie au service du sens ? (ALOES)*. Lyon. 31 March-01 April 2023.

MOORE MAUROUX Susan, 2018, "Phonologie", *in* Françoise Bort, *Rapport du jury de l'agrégation externe d'anglais, session 2018*, p. 70-86.

MOORE MAUROUX Susan, 2022, "English Pronunciation and the Spelling-Sound Code: What Priorities for teachers of EFL?", *in* Monika Pukli, *Pronunciation Matters: Current Perspectives on Teaching and Learning L2 Phonology*, *RANAM* n°55/2022, Strasbourg, Presses Universitaires de Strasbourg, p. 175-191.

NILSSON Camilla, NYBERG Jill & STRÖMBERGSSON Sofia, 2021, "How are speech sound disorders perceived among children? A qualitative content analysis of focus group interviews with 10-11-year-old children", *Child Language Teaching and Therapy*, vol. 37(2), p. 163-175. https://doi-org.bnf.idm.oclc.org/10.1177/0265659021995538

OGDEN Richard, 2012, "Making Sense of Outliers", *Phonetica,* vol. 69(1-2), p. 48-67. DOI: 10.1159/000343197

PY Bernard, 2000, « La construction interactive de la norme comme pratique et comme représentation », *Acquisition et interaction en langue étrangère* [online], vol. 12. DOI: 10.4000/aile.1464

THOMAS Erik R., 2017, *Sociophonetics : An Introduction*, Basingstoke, Palgrave Macmillan.

TRUDGILL Peter, 1972, "Sex, covert prestige and linguistic change in the urban British English of Norwich", *Language in Society*, vol. 1(2), p. 179-195. DOI: 10.1017/S0047404500000488

VAN RIPER William R., 1986, "General American: an ambiguity", in Harold B. Allen & Michael D. Linn (eds.), *Dialect and language variation*, Academic Press, p. 123-135.

WELLS John C., 1982, *Accents of English 1: An Introduction*, Cambridge, Cambridge University Press.

WELLS John C., 2008, *Longman Pronouncing Dictionary*, 3rd edition, Harlow, Longman.

Espaces Linguistiques
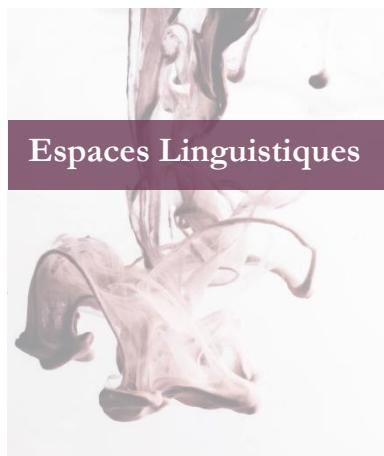
Comment favoriser l'apprentissage des sons de l'anglais par de jeunes locuteurs non natifs ?

How to favour early learning of English sounds by non-native speakers?

Christelle EXARE

Université Sorbonne Nouvelle

PRISMES (EA 4398), Labex EFL

christelle.exare@sorbonne-nouvelle.fr

Jacqueline VAISSIÈRE

Université Sorbonne Nouvelle

LPP (CNRS - UMR 7018), Labex EFL

jacqueline.vaissiere@sorbonne-nouvelle.fr

Résumé : Cet article présente la réalisation d'un cours expérimental de prononciation des voyelles et des consonnes de l'anglais en L2, spécialement créé pour les écoliers français âgés de 5 à 9 ans. Le cours, fondé sur les principes de CleanAccent©, est conçu pour un apprentissage phonétique préliminaire à toute acquisition du vocabulaire, visant à éviter la fossilisation d'erreurs de prononciation couramment observées. Après quelques remarques générales sur l'enseignement de l'anglais en France, cet article décrit le but de l'expérience : dix heures offertes en présentiel dans une classe pour un apprentissage complet de tous les phonèmes. Il présente ensuite en détail les stratégies pédagogiques sélectionnées durant ces séances d'entraînement : a) la progression dans les 10 leçons guidée par les oppositions phonologiques à acquérir, b) la représentation de chaque voyelle de l'anglais (API) par une image et par une séquence de gestes, avec l'évitement du recours à l'orthographe, c) un jeu de 20 planches de loto illustrant du vocabulaire répondant à la progression choisie et accessible aux enfants, d) l'utilisation de plusieurs locuteurs pour favoriser la phonologisation, e) des vidéos d'accompagnement sur YouTube, f) des jeux mis en ligne, g) et l'apprentissage passif éveillé. Les ressources, réelles et virtuelles, ont été mises à l'essai en partie dans une école d'Évreux. L'expérimentation mène à de nouveaux projets que nous décrivons à la fin de cet article.
Mots clés : Anglais L2, écoles françaises, apprentissage précoce, phonétique, phonologie

Abstract: This paper presents an experimental pronunciation course on vowels and consonants, specially made for French schoolchildren aged 5 to 9. The course, based on the principles of CleanAccent©, is designed to avoid the fossilization of commonly observed errors, through a phonetic learning program preceding any vocabulary acquisition. After some general remarks on L2 English teaching in France, this article describes the goal of the experiment: offering 10 teaching sessions covering all the phonemes of English in a classroom setting. Then the choices that were made are presented: a) lesson mapping in ten steps—targeting phonological contrasts, b) an illustrated IPA vowel chart with vowel representation through pictures and gestures, spelling being altogether ruled out, c) a set of 20 lotto boards showing child-friendly vocabulary and matching the lesson plans, d) multiple-speaker auditory inputs to favor phonologization, e) companion videos to the course available online, f) online games, g) passive learning through relaxation exercises. Our real and virtual resource pack was tested in a school in Evreux. The experiment has opened up new projects that are described at the end of this article.
Keywords: L2 English, French schools and homes, early pronunciation teaching, phonetics, phonology

Introduction

% niveau faible

Le faible niveau des élèves francophones dans leur apprentissage de l'anglais est bien connu. La première enquête européenne de 2012 (*Première Enquête européenne sur les compétences linguistiques* 2012) a évalué les compétences en écoute, en écriture et en lecture des étudiants de 14 pays européens et a clairement montré que les étudiants français sont peu performants. Malheureusement, nous manquons de données officielles actualisées et détaillées concernant le niveau des élèves français en anglais oral L2. En effet, CEDRE 2004 et 2010 (*Méthodologie du Cycle des Évaluations Disciplinaires Réalisées sur Échantillon (Cedre) en fin d'école et fin de collège*, 2004) concerne toutes les compétences langagières sauf l'expression orale. Cette dernière est cependant intégrée aux évaluations à partir de 2016. Le CNESCO (Conseil National d'Évaluation du système SCOlaire) observe qu'au collège, en anglais, « 75 % des élèves ne sont pas capables de produire une langue globalement correcte », avec « une prononciation hasardeuse » (CNESCO, 2019, p. 18). Ce diagnostic médiocre de la situation de l'apprentissage de l'anglais en France est partagé par la communauté scientifique et fait l'objet de nombreux rapports (Legendre, 1995, 2003 ; Halimi, 2012 ; L'état de l'École : 32 indicateurs sur le système éducatif français, 2012 ; Beuzon, Boucé, Garcia, Keskpaik & Marchois, 2013) et propositions (Narcy-Combes, Tardieu, Le Bihan, Aden, Delassalle, Larreya & Raby, 2008 ; Taylor & Manes-Bonnisseau, 2018).

% difficulté de la prononciation

La prononciation de l'anglais est difficile pour les apprenants francophones, pour de nombreuses raisons bien identifiées (Vuletić, 1965 ; Adamczewski & Keen, 1973 ; Faure, 1975 ; Huart, 2002 ; Capliez, 2011). Les problèmes se situent aux niveaux segmental (Delattre, 1965 ; Exare, 2017, entre autres) et suprasegmental (Tortel, 2009 ; Horgues, 2010, entre autres). De nombreux facteurs pourraient expliquer les difficultés des élèves francophones dans leur prononciation de l'anglais L2 : les *dispositions articulatoires* (Honikman, 1964) et les *modes phonétiques* différents (Delattre, 1953), les *inventaires phonémiques disjoints* (Ginésy, 2005), la *surdité accentuelle* à laquelle sont exposés les Français (Dupoux, Pallier, Sebastian & Mehler, 1997) ou encore une *pédagogie inadaptée* (Herry-Bénit, 2008) abordant insuffisamment la motivation et l'évaluation (Herry-Benit & Pillot-Loiseau, 2012). Ceci est regrettable, car on sait, par exemple, que l'entraînement phonétique atténue la surdité accentuelle (Carpenter, 2015), améliore la catégorisation des phonèmes (Iverson, Pinet & Evans, 2012) et les réalisations phonétiques des mots commençant par /h/ ou une voyelle (Exare, 2022). On recommande aussi un apprentissage précoce (Gaonac'h & Macaire, 2019), censé être un fort vecteur de motivation (Voise, 2018).

1. Contexte d'enseignement de l'anglais dans les écoles en France en 2022

1.1. Quel public ?

En France, l'anglais est théoriquement enseigné dès la scolarité en école élémentaire. Les élèves de 6-7 ans sont censés être initiés à une langue étrangère par un cours hebdomadaire de 90 minutes (*Loi n° 2013-595 du 8 juillet 2013 d'orientation et de programmation pour la refondation de l'École de la République*, 2013). Toutefois, cette politique ambitieuse française peine à être mise en œuvre de manière adéquate pour deux raisons principales. Premièrement, les enseignants du primaire manquent de formation (Voise, 2010)[17] pour enseigner sans intermédiaire la deuxième langue que les écoles et les familles privilégient parmi toutes les autres pour l'enseignement : l'anglais (Halimi, 2012, p. 20). Deuxièmement, les mathématiques et le français ont longtemps eu la priorité sur les langues secondes dans notre système scolaire national français[18]. Or, l'anglais est appris plus tard, au collège et au lycée, par 97 % des élèves dans notre système français (données de 2020)[19].

1.2. Qu'enseigne-t-on ?

Dans les écoles françaises (Capliez, 2015) **tout comme à l'étranger** (Frost & Henderson, 2013 ; Pourhosein Gilakjani & Sabouri, 2016)**, la prononciation de l'anglais n'est pas toujours mise au premier plan.** Du primaire au secondaire, l'anglais est enseigné par l'entrée civilisationnelle ou culturelle[20], et selon les principes du *Cadre européen commun de référence pour les langues (2018)*[21] qui encourage l'intelligibilité plutôt que la recherche de l'exactitude d'un modèle natif en matière de prononciation. Par exemple, au cycle 3, la figure 1 montre que l'enseignement s'articule autour de quatre activités langagières (expression orale en continu, expression orale en interaction, compréhension auditive, expression écrite).

---

[17] Malgré les dispositions prises par les gouvernements ces dernières années pour 1) former les professeurs des écoles avec l'introduction de la langue vivante étrangère comme épreuve orale facultative d'admission au concours de professeur des écoles depuis 2020 et 2) les motiver, par exemple avec l'ouverture des certifications complémentaires aux enseignants du premier degré à partir de 2019 (cf. aptitude à enseigner une discipline non linguistique en langue vivante).
https://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=143919

[18] Evaluations repères CP, CE1 (cours préparatoire et cours élémentaire 1ère année) Document de travail n° 2021-E06. Série Études, Novembre 2021. Ministère de l'Éducation nationale, de la jeunesse et des sports. Direction de l'Évaluation, de la Prospective et de la Performance (DEPP).

[19] https://www.education.gouv.fr/renforcement-de-l-apprentissage-de-l-anglais-oral-au-college-et-au-lycee-4931

[20] https://eduscol.education.fr/1726/programmes-et-ressources-en-langues-vivantes-voie-gt ;
https://www.education.gouv.fr/les-programmes-du-college-3203

[21] https://www.coe.int/fr/web/common-european-framework-reference-languages

Figure 1. Programme du cycle 3 de consolidation (CM1, CM2, 6ème) en langues vivantes22

Langues vivantes

Au cycle 3, l'enseignement de la langue vivante étrangère ou régionale vise l'acquisition de compétences et de connaissances qui permettent l'usage plus assuré et plus efficace d'une langue autre que la langue française. Des situations de communication adaptées à l'âge, aux capacités cognitives, aux intérêts des élèves, contribuent à la construction de connaissances langagières, permettant d'atteindre le niveau A1 du Cadre européen commun de référence pour les langues (CECRL) dans les cinq activités langagières.

Compétences travaillées :

— Écouter et comprendre
— Parler en continu
— Écrire
— Réagir et dialoguer
— Découvrir les aspects culturels d'une langue vivante étrangère et régionale

Au programme :

Les thématiques ou les types de supports (théâtre, cinéma, poésie...) mentionnés en cycle 2 peuvent être repris en veillant à proposer une progression sur l'ensemble de la scolarité obligatoire et en évitant les redondances, l'objectif d'enrichissement linguistique restant lié aux autres enseignements dispensés. Au cycle 3, les connaissances culturelles sont réparties selon trois axes :

— La personne et la vie quotidienne
— Des repères géographiques, historiques et culturels dans la langue étudiée
— L'imaginaire

Source : https://www.education.gouv.fr/les-programmes-du-college-3203

La langue étrangère est devenue le vecteur socioéconomique de la mondialisation. En conséquence, les apprenants non natifs, en tant que futurs agents économiques et sociaux, sont encouragés à apprendre la langue étrangère par des approches fondées *sur des tâches* (Manoïlov, 2019), imitant la réalité du marché de l'emploi et des échanges internationaux. La fluidité dans les interactions et la capacité à se faire comprendre sont perçues comme étant plus utiles que l'exactitude grammaticale, lexicale ou phonologique. L'objectif principal est qu'un locuteur comprenne et soit compris. La langue étrangère sert à interagir avec d'autres et à agir sur son environnement et ses pairs. L'essentiel est alors de développer des compétences sociales et pragmatiques en langue étrangère, et non pas d'avoir une « bonne » prononciation. Notons qu'il est difficile de définir exactement ce qu'est une bonne prononciation. Une « bonne » prononciation ou une prononciation suffisamment bonne est, selon nous, une prononciation qui, si elle n'est pas strictement identique phonétiquement à celle d'un natif, évite les glissements sémantiques grâce à

---

22 En juin 2023, le nouveau programme de cycle 3 intègre une compétence supplémentaire en langues vivantes : « Lire et comprendre ». https://eduscol.education.fr/document/50990/download. De plus, le Ministère de l'Éducation Nationale publie un « Guide pour l'éveil à la diversité linguistique en maternelle » qui, tout en renouant avec l'objectif ancien d'éveiller les jeunes enfants du cycle 1 à la pluralité des langues du monde, laisse au lecteur pédagogue la responsabilité de créer les ressources pédagogiques et didactiques lui-même.
https://eduscol.education.fr/document/50921/download?attachment

la conscience des oppositions phonologiques de la langue (contrastes entre phonèmes) et grâce à l'existence des habitudes articulatoires, rythmiques (répétition de schémas dans le temps) et mélodiques (variation et courbe de la fréquence du fondamental) de la langue cible.

1.3. Comment l'anglais est-il enseigné dans les petites classes ?

Le ministère français de l'Éducation nationale publie des « programmes articulés au socle commun de connaissances, de compétences et de culture » (*Socle commun de connaissances, de compétences et de culture*, 2015) avec des ressources d'accompagnement23, guides24 et circulaires25 pour l'enseignement de l'anglais L2 dans les petites classes. Deux observations doivent être faites. Premièrement, les progressions suggérées reposent peu sur des bases phonétiques et phonologiques. Deuxièmement, les consignes restent des suggestions, des pistes et des exemples, sans qu'aucune directive ni méthode d'enseignement officielle, en version papier, ne soit publiée par le Ministère. Les éditeurs, quant à eux, investissent peu ce domaine. Les enseignants du primaire en formation sont invités à concevoir leur propre matériel. Pour ce faire, ils bénéficient d'un grand nombre de ressources en ligne pour concevoir leurs plans de cours26. Ces ressources, par ailleurs très utiles, ont des inconvénients : i) elles sont éparpillées sur internet, de sorte que leur consultation est chronophage27, ii) elles sont parfois payantes28, iii) elles sont surtout fondées sur des principes communicatifs et culturels et offrent peu d'éclairage sur la prononciation. Finalement, les ressources de type « phonics » (EEF, 2018)29 traitent rarement de l'articulation de la parole proprement dite, mais plutôt de l'orthographe. Or, on sait que l'absence de correspondance stricte graphie-phonie en anglais et en français rend difficile l'apprentissage de la lecture et de l'écriture pour les élèves (Moll, Ramus, Bartling *et al.*, 2014). Notre positionnement est donc d'éviter le

---

[23] https://eduscol.education.fr/164/langues-vivantes-cycles-2-3-et-4

[24] https://eduscol.education.fr/159/guide-pour-l-enseignement-des-langues-vivantes-etrangeres

[25] https://www.education.gouv.fr/bo/19/Hebdo22/MENE1915455N.htm
À l'école maternelle, l'apprentissage précoce des langues est optimisé lorsque les enfants jouent, réfléchissent, s'exercent et entraînent leurs capacités mnésiques.

[26] Edumoov est connu des enseignants stagiaires de l'enseignement élémentaire https://www.edumoov.com/
Citons aussi La clé des langues : https://cle.ens-lyon.fr/
Les sites académiques : https://primlangues.education.fr/references/sites-academiques
Neopass action : http://neo.ens-lyon.fr/neo
Captain Kelly (2022) : https://eduscol.education.fr/2974/enseigner-l-anglais-l-ecole-avec-captain-kelly

[27] Voir la circulaire de 2019 pour les LV en école maternelle : https://eduscol.education.fr/2326/langues-vivantes

[28] https://www.lasalledesmaitres.com/ressources/?swoof=1&product_tag=anglais

[29] Education Endowment Foundation-Evidence Summaries :
https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/phonics/

recours à l'orthographe pour l'apprentissage de la prononciation de l'anglais dans les petites classes, c'est-à-dire en cycle 1 ou 230.

1.4. Quelle est la situation institutionnelle aujourd'hui ?

Le partenariat entre enseignants et chercheurs a longtemps été difficile (Desgagné, 1997). Le contexte institutionnel semble dorénavant favorable pour que des expériences de recherche collaborative comme la nôtre soient régulièrement menées dans de vraies classes (Article L. 314-2, Loi pour une École de la confiance, 2019). De plus, la coéducation est devenue une priorité du gouvernement français31, **les relations entre les familles et les écoles étant complexes et tendues depuis longtemps** (Meirieu & Hameline, 2000). Ce projet a l'ambition de favoriser un triple soutien de l'enfant-apprenant : les maîtres et maîtresses, les familles, et nous-mêmes, enseignants-chercheurs. Les recommandations gouvernementales suggèrent de privilégier des séances d'exposition courtes et régulières des enfants à l'anglais. Il est conseillé 15 à 20 minutes par jour à l'école primaire (Taylor & Manes-Bonnisseau, 2018 ; *Un plan d'actions pour une meilleure maîtrise des langues vivantes étrangères,* 2022).

2. Les principes de base de CleanAccent©

À partir de ces observations, nous avons conçu une méthode modulable, évolutive et ludique. Inspiré par les résultats récents de la recherche en psycholinguistique sur la mémorisation et l'apprentissage passif et actif, l'objectif principal de notre méthode est d'éviter la fossilisation des erreurs de prononciation dans l'apprentissage précoce de l'anglais. La première étape de la méthode CleanAccent© est une pratique intensive de la production et de la perception des sons de l'anglais L2, principalement dans des mots monosyllabiques, avec une approche de l'accent lexical et de la réduction vocalique dans quelques mots dissyllabiques. Les principes de base qui ont été entièrement mis en œuvre sont les suivants :

– Les 20 voyelles et 24 consonnes de l'anglais britannique sont enseignées avec progressivité dans différents contextes : à l'isolé pour les voyelles, dans des mots de préférence monosyllabiques, dans des groupes nominaux partageant la même voyelle (tel que « green sheep », et l'image d'un mouton vert représente la voyelle /i :/ sur le triangle vocalique).

– Il est plus facile d'apprendre des *sons nouveaux* que des sons similaires (Flege, 1987). La progression prend acte des *sons anglais les plus confondus* avec le français, sur la base de (Delattre,

---

1965) et (Gimson, 2001), et des résultats sur les difficultés de perception de certains phonèmes (Flege, 1984 ; Hallé, Best & Levitt, 1999 ; Iverson, Pinet & Evans, 2012).

– Les *gestes articulatoires* nouveaux (comme l'aspiration des trois occlusives sourdes, la diphtongaison en anglais, une moindre anticipation vocalique, et l'effort articulatoire plus important mis sur la syllabe accentuée sur mot).

– Les *contrastes phonologiques* (le recours le plus systématique possible aux paires minimales contrastant des sons proches) sont enseignés (Levis & Cortes, 1995 ; Tuan, 2010 ; Haghighi & Rahimy, 2017).

– *L'association de chaque voyelle de l'anglais et du schéma accentuel des mots multisyllabiques avec des mouvements des mains* est utile durant la phase d'apprentissage et s'apparente pour les enfants à un jeu. Notre approche est explicite (Rosenshine, 1986) et multimodale (Jewitt, 2012). Les correspondances entre activités corporelles et acquisition de la prosodie et des sons d'une langue sont un objet d'intérêt certain et croissant pour la communauté scientifique actuellement (Potapushkina-Delfosse, 2014 ; Chan, 2018).

– Les apprenants sont exposés aux voix de *différents locuteurs* (2 hommes et 2 femmes au minimum) pour favoriser la catégorisation phonologique (Rost & McMurray, 2009 ; Thomson, 2018 ; Nagle & Baese-Berk, 2022, p. 591).

– L'apprentissage implicite du *triangle vocalique* sous forme d'images, une par phonème, est utile (Arleo, 1993). La figure 7 sert de référence constante. L'enfant doit pointer sur l'image représentant la voyelle du mot monosyllabique qu'il entend.

– Le vocabulaire est présenté sous forme d'un *jeu de loto*. Ce jeu de société favorise des échanges verbaux simples entre enfants et enseignants. Les règles du jeu peuvent varier à loisir pour ne pas lasser les enfants. Ils peuvent par exemple rassembler les cartelettes concernant les animaux ou les aliments entendus, jouer à un jeu de loto traditionnel avec une planche à compléter en un temps minimal, s'amuser avec des jeux de devinettes etc. Les enfants débutant l'anglais ont ainsi accès à une interaction aisée sous forme de mots simples dans un apprentissage actif et ludique.

– *L'apprentissage passif* fait partie du programme de révision après la classe. L'exposition passive aux ressources enregistrées et mises à disposition est assurée de façon quotidienne, pendant les phases de détente en classe et à la maison dans l'espace familial.

– Le recours aux technologies numériques avec *l'accès à un site internet* spécifiquement dédié récrée, en dehors de la salle de classe, les conditions d'apprentissage favorables à l'apprentissage de la langue non native (Frost, 2022).

– La variété choisie pour les leçons en classe est *l'anglais britannique standard* pour deux raisons. Tout d'abord, l'anglais britannique est encore, par rapport à l'anglais américain, la variété la plus enseignée en France à l'heure actuelle (Michari, 2023). De plus, les deux enseignantes filmées pour l'expérimentation maîtrisent l'anglais britannique standard.

Concernant les consonnes, on introduit tôt — dès la première et la deuxième leçon — de nouveaux sons : les occlusives sourdes aspirées [pʰ], [tʰ], [kʰ] et la fricative glottale sourde /h/. Une feuille de papier placée par chaque élève devant sa bouche rend l'apprentissage de l'aspiration concret et effectif. Notons que la grande majorité des mots choisis ne commençant pas par une consonne aspirée ou /h/ commencent par des consonnes similaires (sauf /r/).

Concernant les voyelles, la première priorité est de les regrouper en fonction de leur position dans l'espace vocalique: « Difficulty is most predictable in those areas where vowels are closest within the vowel space; thus confusions are very likely within any of the following groups /iː, ɪ/, /e, æ, ʌ/, /ɒ, ɑː, ɔː/, /uː, ʊ/ » (Gimson, 2001, p. 103). L'intégration précoce des diphtongues constitue une seconde priorité pour le choix de la progression (le français standard ne connaît aucune diphtongue). Les diphtongues sont regroupées selon la voyelle relâchée correspondant à leur premier élément (« onglide »). Par exemple, la leçon 2 porte sur /e/, /eɪ/ et /eə/. La leçon 3 traite de /æ/, /aɪ/, aʊ/, etc. (voir Tableau 1).

Deux remarques : Il ne nous a pas été possible de suivre le conseil d'une pratique de l'anglais de 10 à 20 minutes par jour, et les 10 cours proposés ont été étalés sur 10 semaines, avec une proposition d'écoute journalière de 5 à 10 minutes pour répéter et fixer les acquis, disponible sur YouTube.

3. Organisation et choix pour l'École d'Évreux

L'école partenaire est l'École publique élémentaire Victor-Hugo. Cette école fait partie de l'académie de Normandie (zone B). Elle scolarisait 121 élèves au moment de l'expérimentation. Évreux est une ville moyenne d'environ 50 000 habitants, préfecture du département de l'Eure (27). Tout en restant une ville de province, elle devient très attractive pour les navetteurs travaillant sur Paris. Cette école a été choisie parce qu'elle se situait à proximité du domicile de l'une des expérimentatrices. Le quartier est un quartier de centre-ville plutôt favorisé. Le contact a été établi avec le directeur et les enseignantes dès 2020. Le projet a été concrétisé après avoir été doté d'une enveloppe de 5 000 euros dans le cadre du LabEx « Empirical Foundations of Linguistics » (Actions Défis Sociétaux, 2020).

La maîtresse de l'unique classe de CP de l'école est titulaire d'un DEA de chimie. Nous avons été très bien accueillies par l'école.

La classe était composée de 23 élèves (12 filles et 11 garçons), réunis en 4 groupes de 5 ou 6. La maîtresse enseignait à 3 groupes pendant que nous étions dans la salle d'étude contiguë avec un groupe pendant 30 minutes, un après-midi par semaine, pendant 10 semaines. Nous portions tous des masques (enfants et enseignants). Les déplacements dans la salle étaient impossibles et nous avons dû travailler en respectant les mesures d'hygiène liées au COVID.

Notre expérience pilote à Évreux avec quatre groupes de 6 enfants dans 10 sessions hebdomadaires de 30 minutes chacune montre que :

i. 30 minutes est une durée trop courte pour un apprentissage intensif des sons de l'anglais et laisse peu de temps pour la répétition,

ii. un effectif de 6 élèves est peu porteur pour obtenir une bonne dynamique de groupe, notamment en cas d'absence d'un ou deux élèves,

iii. l'exposition quotidienne recommandée est absolument nécessaire pour compléter le cours intensif. L'apprentissage passif a été peu suivi, d'après le retour des enfants eux-mêmes. Le manque de participation de certains parents, pour des raisons diverses et complexes à résoudre, a posé problème pour la progression de leur enfant.

Figure 2. Répartition des tâches. L'élève, qui est « au centre de son apprentissage » (Boutin, 2004, p. 16), est accompagné de trois soutiens. Son professeur (1) accueille les intervenantes phonéticiennes anglicistes (2) qui font les cours. Sa famille (3) participe en lui faisant écouter et réviser les leçons en ligne.
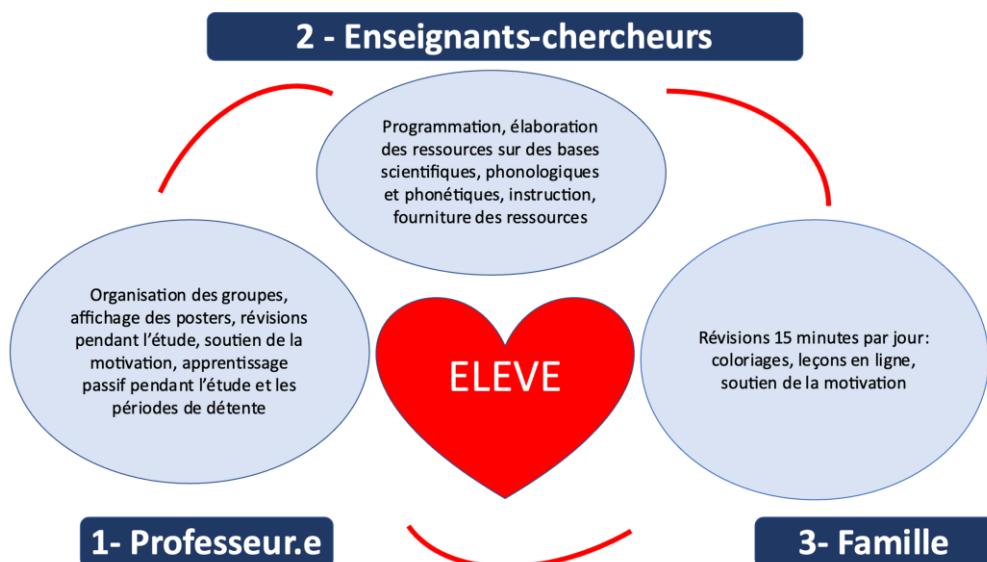
Pour réaliser notre méthode, nous avons procédé à des choix concernant : le vocabulaire (quels mots sélectionner ?), la conception des leçons (quels sons pour chaque leçon, quel ordre de succession dans les leçons ?), le choix des illustrations libres de droit, la nature de l'*input* sonore, la gestuelle associée à chaque phonème, l'illustration de l'API en images et la création de ressources hypnopédiques pour aider les révisions par apprentissage passif. Les supports étaient, pour l'essentiel, prêts avant l'expérimentation à Évreux. Les réactions et commentaires des enfants ont été pris en compte, de semaine en semaine, pour affiner nos choix et notre méthode. Chaque séance (partiellement filmée par la deuxième expérimentatrice) donnait lieu à une concertation collective (deux expérimentatrices, maîtresse, assistante chargée de l'étude du soir) et réflexive (aspects positifs à conserver, aspects négatifs à améliorer, causes des événements, conséquences à prévenir) visant à améliorer nos choix d'enseignement et nos pratiques pédagogiques.

3.1. Vocabulaire

Le choix du vocabulaire est un choix difficile, parce que nous sommes évidemment limitées par la langue elle-même ! Nous avons essayé de faire le meilleur compromis possible entre les contraintes suivantes :

– Le vocabulaire sélectionné décrit des choses concrètes que les enfants apprécient, en particulier la nourriture et les animaux.

– Les mots lexicaux monosyllabiques sont privilégiés pour l'apprentissage des segments, car ils présentent souvent des paires minimales. Par exemple, « sheep » s'oppose à « ship », « cat » à « kite » (Figure 3).

Figure 3. Paires minimales CV ou CVC avec V = /æ, aʊ, aɪ/
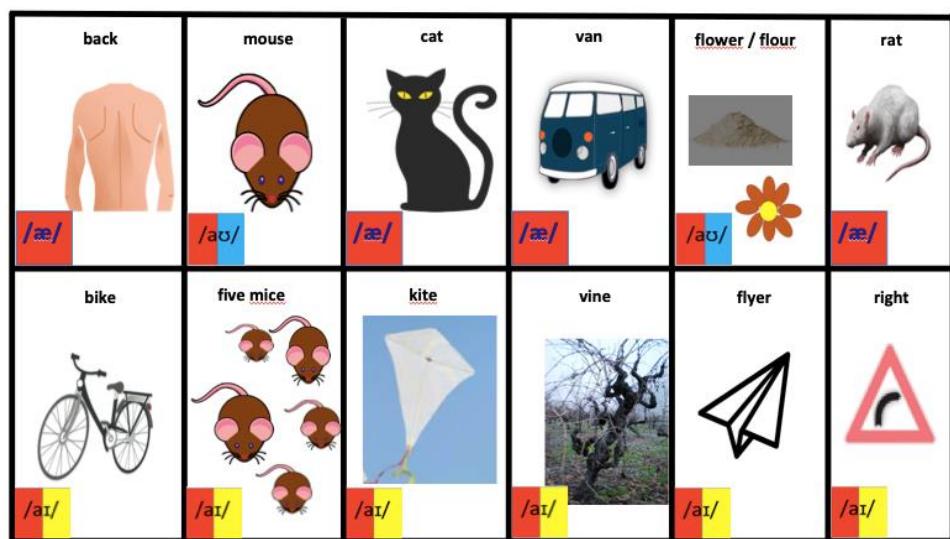
– Des mots plus longs de 2 ou 3 syllabes sont utilisés pour une première approche de l'accent lexical et pour faire prendre conscience de la réduction des voyelles. Citons par exemple « tomato, hippo, banana, potato ».

– Des mots d'emprunt ou des mots similaires sont utilisés pour mettre en valeur la différence de prononciation entre le mot français et le mot anglais, tout en réduisant la difficulté d'accès au sens (Rogers, Webb & Nakata, 2015). Citons par exemple « rat, toast, tennis, carrot, tiger », etc.

– Les mots sont enseignés en paires ayant le même son cible pour faciliter la mémorisation, en faisant varier les environnements phonémiques et en entraînant les élèves à apprendre les mots en contexte et dans des collocations (Pawley & Syder, 1983 ; Nattinger & DeCarrico, 1992 ; Schmitt, 2004 ; Boer & Lindstromberg, 2005)32. Le triangle vocalique (Figure 6) présente des syntagmes (surtout nominaux) : « a green sheep, a little pig, a grey train, a red hen, a bear with a pear ». Le nombre d'images est limité pour certains phonèmes, comme /ʊə/, en raison de l'absence d'images concrètes pour les illustrer.

– Les mots sont présentés sur des cartelettes réunies dans des planches permettant des jeux de loto, de devinettes etc. L'intérêt des mémo-cartes, d'une part, et de l'apprentissage des mots à l'isolé, d'autre part, bien qu'apparemment en contradiction avec les méthodes communicatives, semble très utile pour la communication entre très jeunes enfants (Hilton, 2022, p. 140-144).

– Chaque élève a appris à prononcer son nom anglicisé ainsi que celui de ses camarades.

---

32 Par exemple, Bouvet (2021, p. 175), à la suite de Chen-Géré et al. (2000), propose de faire produire /ʊ/ avec « The Jungle book is a good book! ».

3.2. Progression

Le tableau ci-dessous décrit le déroulé des 10 leçons pour l'enseignement des phonèmes de l'anglais à des apprenants francophones débutants.

Tableau 1. Déroulé des 10 leçons pour l'enseignement des phonèmes de l'anglais à des apprenants francophones débutants

Phonèmes	Titre de la leçon	16 mots	8 mots	Mots en contexte
/p, t, k, h/	Pat's pet Hello! Kate's kite Hi!	Hello, hi, hippo, hotdog, potato, panda, pizza, Pat, toast, tennis, tiger, tomato, camera, caramel, karate, carrot	Panda, pizza, tiger, tennis, carrot, karate, hotdog, hippo	
/e, eɪ, eə/	Hey! let's play a game! Hey! Let's paint a pear!	Pear, parents, tennis, train, red, grey, hen, hair, air, pepper, play, friend, wear, square, paint, paper	Hen, red, pear, bear, grey, train, hair, air	Grey train, red hen. The bear has a pear.
/æ, aɪ, aʊ/	Hi Dad! Bye bye my cat!	Cat, kite, hat, rat, mouse, mice, right, flower, ouch, black, white, brown, fly, flyer, sky, cow	Cat, hat, black, white, kite, cow, mouse, brown	Black cat, brown cow, white kite
/ɔɪ, ɒ/	What an odd toy in this orange box! Boing !	Box, coffee, dog, doll, donkey, fox, frog, waffle, boy, coin, coy, enjoy, foil, noisy, toy, turquoise	Dog, fox, box, frog, doll, boy, toy, noisy	Orange fox, odd frog, noisy toy, coy boy. A doll in a box.
/ə, əʊ/	Hello! No, no, no, no!	Toast, nose, rose, goat, soda, boat, yellow, old, banana, gorilla, pizza, sugar, river, water, sister, brother	Yellow, rose, toast, goat, banana, gorilla, water, river	Yellow rose. The gorilla has a banana.
/ʌ, ɜː/	Yummy Mum, Yum yum cucumbers, Yuck Yuck burgers and gherkins	Bird, turtle, girl, turkey, earth, dessert, dirty, purple, bus, duck, brush, cup, plum, monkey, funny, young	Girl, purple, bird, turtle, bus, duck, monkey, funny	Purple bird, funny monkey. A duck on a bus
/ɪ, ɪə, iː/	Easy peasy itsy bitsy	Ship, sheep, big, little, bee, beer, pink, green, kiss, pierce, peace, slip, sleep, pig, king, queen	Ship, pink, pig, big, beer, green, bee, sheep	Green sheep, big ship, pink pig, green tea, green pea
/ɑː, ɔː/	A horse at the door and a car in the garden!	Dance, garden, car, castle, park, bar, heart, arm, ball, horse, door, boar, pork, draw, small, tall	Horse, ball, small, tall, car, park, heart, star	Small horse, a car in the park
/ʊ, uː, ʊə/	Good cook !	Cook, pull, look, poor, blue, pool, Luke, tooth, moon, scooter, goose, book, wolf, curious, furious, sure	Blue, moon, pool, book, pull, cook, wolf, poor	The blue moon, a good cook
Les consonnes	/ʃ/ like sheep, /ʧ/ like cheap	Rice, lice, sheep, cheap, bus, buzz, led, red, sink, think, ship, chip, fan, van, tree, three	Cheap, jeep, think, sink, rice, lice, jeans, gin	The cheap Jeep

À chaque leçon correspondent des planches de loto. Les images de Cleanaccent© sont libres de droits. La plupart ont été récupérées sur Pixabay. De nombreuses images ont été transformées et adaptées manuellement pour répondre à nos besoins.

Figure 4. Planche pour la voyelle /ɑ:/

3.3. Déroulement d'une leçon

Une leçon se concentre sur 2 ou 3 phonèmes. Chaque leçon s'appuie sur 2 planches de 8 mots ou plus, avec les images et sons associés, les gestes correspondants, une image positionnant le phonème dans le triangle vocalique, des exercices et jeux pour réactiver les connaissances à la maison. Le français — ou la langue maternelle des apprenants — peut être utilisé pour lever les ambiguïtés sémantiques potentielles (Littlewood & Yu, 2011).

Après les salutations et la prise de renseignements autour des devoirs faits (coloriages, écoute des fichiers en ligne, etc.), la leçon commence par une description de l'objet de la leçon : l'articulation, accompagnée des gestes, de phonèmes différents dans des mots « simples » (transparents, courts, participant d'une thématique aimée des enfants), comme par exemple : « Hi ! » comparé à « aïe », « Hello », « toast », « panda », « carrot ». Ces mots-ci, enseignés à l'étape 1, permettent d'introduire l'aspiration phonétique des occlusives sourdes (un petit souffle doit faire trembler la feuille) en parallèle du son /h/ inconnu de l'inventaire du français ainsi que le /r/ apical. Dès la première leçon, les enfants sont encouragés à porter leur attention sur leurs articulateurs (la glotte et la langue notamment). Ils sont amenés à expérimenter et sentir la formation de la voix dans leur corps et non seulement dans leur cerveau. Le port du masque étant obligatoire, des visières transparentes couvrant le bas du visage des expérimentatrices sont utilisées pour cette présentation.

Après cette phase liminaire se concentrant uniquement sur la perception des phonèmes et de leurs caractéristiques phonétiques dans des mots courts et isolés, les enfants (un par un, puis en chœur) sont invités à reproduire les sons enseignés à l'isolé. Après la présentation des phonèmes, les enfants apprennent les 16 mots des deux planches dont ils disposent devant eux. Les mots, présents

sur les cartelettes individuelles composant les planches, sont aussi projetés sur un mur blanc grâce à un vidéoprojecteur. On s'assure dans un premier temps que les référentiels des images sont bien identifiés en français. Puis les enfants prononcent les mots anglais après le modèle (l'expérimentatrice ou une autre voix enregistrée et disponible sur le site privé en ligne). Certains enfants accompagnent spontanément leurs productions des gestes enseignés au préalable.

En cas de fatigue ou d'excitation extrême des enfants, le vocabulaire est diffusé grâce aux fichiers d'apprentissage passif éveillé.

Plusieurs fois par séance (début, fin ou milieu) de petits tests (évaluations intermédiaires) fréquents et donnés à un rythme tonique permettent de consolider les acquis (Schwieren, Barenberg & Dutke, 2017), tout en stimulant les élèves et leur plaisir de relever des défis. En production, il s'agit par exemple pour l'enseignant de projeter 10 images en demandant aux enfants de nommer les objets en anglais le plus rapidement possible. Il peut être demandé aux élèves de se questionner entre eux, en montrant une image. En réception, les enfants aiment compléter leur grille de loto le plus vite possible, ou pointer du doigt les images correspondant au mot entendu sur des tableaux montrant tous les objets de la leçon rassemblés dans un même paysage bucolique (Figure 5).

Figure 5. Planche de révision des mots appris durant les trois premières leçons

À ce niveau élémentaire, et parce que notre objectif est phonético-phonologique (mises en place des sons, de l'accent lexical et de la réduction vocalique), tout le travail est mené au niveau du mot ou du syntagme nominal.

La séance se termine par l'explication des révisions à faire à la maison. Les consignes sont données oralement aux enfants et à l'écrit à l'attention des parents (avec un bref message et le lien vers le site privé). La transmission d'information peut être dupliquée grâce à l'intermédiaire de la maîtresse (par mail, de visu).

Des renforçateurs et des objets personnalisés sont utilisés pour soutenir la motivation des enfants. Chaque élève a un classeur qui lui est propre, et avec un chevalet sur lequel est inscrit son prénom anglicisé et transcrit en API, avec un espace pour colorier et dessiner son autoportrait ou un animal fétiche.

Figure 6. Les supports offerts à la première leçon

Quand c'est possible, des objets sont apportés comme référentiels complétant les images (café, gaufre, tomate, banane, pomme, carotte, panda en peluche). Les rétroactions verbales sont systématiques, avec au moins une félicitation en cas de production satisfaisante. En cas de production insatisfaisante, un recours aux pairs, une demande de répétition après le modèle, accompagnée d'un geste et éventuellement des précisions articulatoires sont fournies par l'enseignante. Nous nous imposons une négociation de la forme et du sens (Hendrickson, 1978 ; Lyster & Ranta, 1997 ; Lyster, 1998) jusqu'à ce qu'une réalisation acceptable soit perçue (le glissement sémantique est évité, les sons bien catégorisés, l'accent de mot bien placé et l'intonation appropriée). Le renforcement positif optionnel (Wheatley, West, Charlton *et al.*, 2009 ; Diedrich, 2010) consiste en la distribution d'autocollants à placer sur une feuille d'évaluation formative et positive personnalisée (« reward chart »).

3.4. Ressources d'accompagnement

Les ressources d'accompagnement sont de deux types : papier et numérique.

### 3.4.1. Ressources en papier pour la classe

Comme indiqué plus haut, on confie aux enfants des pochettes à leur nom anglicisé contenant les planches de lotos, leur fiche d'évaluation qu'ils complètent au fur et à mesure avec les stickers reçus

(« reward chart »), le triangle vocalique et les courriers aux parents, avec le lien vers le site privé. On offre à l'école des affiches de grande taille qui sont collées sur le tableau disponible de la salle d'étude. Parmi ces posters, on trouve :

– Un tableau permettant de classer les mots par champ lexical. Les enfants peuvent jouer à aimanter les mots tout en les répétant pendant leur temps d'étude.

– Un tableau des pronoms personnels, des articles et des adjectifs possessifs les plus courants, avec leur prononciation. Sans être utilisé pour cette phase d'apprentissage des phonèmes, il sera néanmoins utile les années ultérieures, quand les élèves passeront du mot à la phrase.

– Enfin et surtout, l'API sous forme d'images. En effet, l'emploi de symboles de l'API et de mots clés semblent être des stratégies appréciées des apprenants pour améliorer leur prononciation de l'anglais L2 (Mompean & Fouz-González, 2021). Notons que le rapport entre les symboles utilisés pour décrire le système phonologique d'une langue et la correspondance de ces symboles avec la réalité phonétique évolue avec le temps.[33] L'essentiel pour les enfants est d'associer la bonne image (« green sheep ») avec le son d'un nouveau mot (par exemple « read ») et non avec un symbole de l'API.

– Un triangle vocalique (« vowel chart ») sous forme d'images représentant les voyelles de l'anglais britannique standard est présenté dans la figure ci-dessous. Il présente les sons illustrés par des syntagmes comportant deux fois le même son dans des syllabes ouvertes ou fermées. Dans le sens inverse des aiguilles d'une montre, on a: *a green sheep*, *a little pig*, *a grey train*, *a red hen*, *a bear with a pear*, *a white tiger*, *a deer has beer*, *a black cat*, *a good wolf*, *a car in a park*, *an orange fox and a dog*, *a duck on a bus*, *a brown cow*, *a pizza*, *a girl with a turtle*, *a horse and a ball*, *a blue moon*, *a good-looking wolf*, *he's surely poor*, *a yellow rose*, *a boy with a toy*.

---

[33] Par exemple, Lindsey (2019) propose de transcrire /e/ par /ɛ/, tout comme Lilly & Viel (2001). Il observe aussi que les diphtongues centralisantes sont en réalité prononcées comme des monophtongues longues :/ɪə/ et /eə/ seraient prononcées [ɪː] et [ɛː]. Le blog de Geoff Lindsey présente de nombreux exemples montrant l'écart entre la prononciation RP et l'anglais britannique standard actuel (https://www.englishspeechservices.com/). Cependant, en l'absence d'études expérimentales dans ce domaine, nous avons choisi comme repères pour nos transcriptions les deux dictionnaires LPD et CEPD (Wells, 2008 ; Jones, 2011).

Figure 7. Le triangle vocalique illustré de Cleanaccent© pour l'anglais
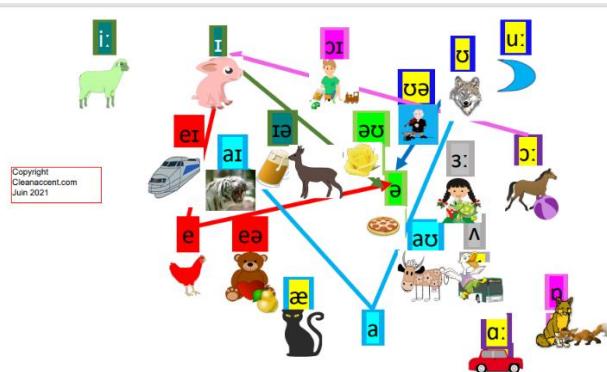
3.4.2. Ressources numériques

Tout le matériel est accessible en ligne sur un espace privé et les vidéos sont dupliquées et doublées sur la chaîne youtube. Elles sont disponibles sur demande. Pour y accéder sur Youtube, il suffit de rechercher « youtube cleanaccent »34.

Le recours à des gestes pédagogiques dans l'enseignement d'une langue étrangère est réputé (Herry-Benit & Cadet, 2014 ; Hilton, 2022, p. 158). Il connaît un regain d'intérêt dans l'idée de la phonologie incarnée (Tellier, 2010 ; Cauna, Tellier & Colé, 2021 ; Bottineau, 2022). Des études expérimentales cherchent à mesurer l'efficacité des gestes pédagogiques accompagnant l'enseignement de la prononciation aux niveau segmental et suprasegmental, mais peu d'effets ont cependant été observés (Alazard-Guiu, Roa, Ferrané, *et al.*, 2022).

Au plan segmental, les vidéos de Maëlle Amand, fondées sur Gattegno (1972), montrent des gestes utiles pour articuler les voyelles. La tension des voyelles tendues apparaît sous la forme d'un élastique qu'on étire, comme dans la figure ci-dessous. Les voyelles relâchées sont associées à un geste tranchant vertical, de bas en haut. Ces vidéos sont disponibles sur Youtube.

---

34 https://www.youtube.com/@cleanaccentcom

Figure 8. Geste de Maëlle Amand. Illustration de /aɪ/

Source : https://cleanaccent.com/british_english/miming-english-vowels-online/

Au plan suprasegmental, nous représentons :

– les syllabes portant l'accent lexical par des gestes d'ouverture de la main vers l'avant

– les syllabes inaccentuées par un repli de la main vers l'arrière, poing serré. La Figure 10 illustre un exemple. Plusieurs schémas sont utilisés pour illustrer le profil prosodique de chaque syllabe dans le mot, en fonction de la position de la syllabe par rapport aux frontières du mot et à l'accent lexical. Le schwa est indiqué de façon particulière.

Figure 10. Le mot « panda » illustré. Dans l'ovale rouge : les gestes des mains pour la succession des syllabes (accentuée puis inaccentuée). Dans l'ovale bleu : le dessin représentant la courbe mélodique. Les chiffres ainsi que la taille des ronds et des ovales indiquent la saillance perceptive des deux syllabes
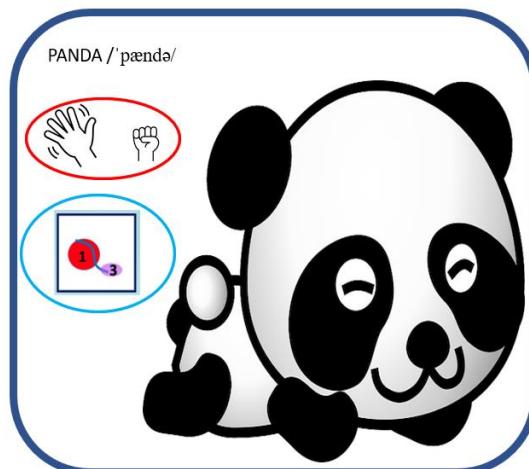
3.4.3. Vidéos de révision pour les leçons et vidéos pour un apprentissage passif

De nombreux chercheurs avancent que l'apprentissage passif éveillé peut avoir des effets bénéfiques sur la mémorisation d'une langue étrangère. Les travaux de l'époque de la Guerre Froide (Hoskovec, 1966 ; Pinsker, 1971) connaissent un regain d'attention (Odendaal, 1987 ; Bancroft, 2005 ; Kapsi, Katsantoni & Drigas, 2020 ; Kapsi, Katsantonis & Drigas, 2021).

Nos films sonorisés accessibles en ligne suivent les principes de l'apprentissage en boucle : les mots sont présentés de façon itérative. Les vidéos peuvent être utilisées dans différents contextes non exclusifs les uns des autres : en début de séance pour réviser ou pour se plonger dans le bain, pour un rebrassage distal et même à tout autre moment de la semaine pour se détendre ou pour se concentrer.

4. Discussion

4.1. Résultats de l'expérience pilote

L'objectif premier de notre expérimentation à Évreux était de vérifier la pertinence de nos choix pour améliorer nos supports cartonnés et numériques. Par exemple, nous avions intégré le mot « oyster » à la liste de mots enseignés dans une phase préliminaire. Devant le dégoût exprimé par les enfants lors de notre expérimentation à Évreux, nous avons décidé de remplacer ce mot par un autre mot.

Pendant la dernière séance d'évaluation, nous avons fait des enregistrements. Malheureusement, les masques, le bruit ambiant et l'absence de plusieurs enfants partis en vacances d'été rendent impossibles des analyses statistiques sur les acquis. Nous avons cependant pu compter le nombre de mots différents bien prononcés, produits spontanément à partir des images. Sur un total de 109 mots enseignés, en 7 séances de 25 minutes par élève, les 18 enfants enregistrés ont restitué, de façon correcte, en moyenne 15 mots chacun. Les 4 enfants qui ont été le plus accompagnés par leurs parents pour des révisions à la maison à partir de nos ressources en ligne (d'après leurs dires et ceux de la maîtresse) ont, eux, produit 26 mots correctement.

Globalement, les enfants n'ont pas produit d'erreurs sémantiques lors de l'interrogation orale finale. Notons cependant que les confusions en production concernent essentiellement des mots commençant par la même consonne : « donkey/duck », « pepper/pear », « queen/cow », « turkey/turtle ». Ceci va dans le sens du modèle psycholinguistique de la cohorte de Marslen-

Wilson (Marslen-Wilson & Welsh, 1978) qui théorise que le lexique est classé par la première consonne du mot dans le lexique mental des auditeurs.

4.2. Coordination avec les familles

Cette première expérience nous a permis de tirer la leçon suivante. **Plus de collaboration avec les familles sera nécessaire** pour encourager les entraînements à domicile qui ont peu été suivis.

Nous avons convenu, lors d'une prochaine expérimentation,

a) de convier les parents au tout début à une réunion informelle pour exposer les objectifs et les méthodes de notre intervention, solliciter les familles en négociant ce à quoi elles peuvent s'engager et distribuer les formulaires de consentement éclairé pour les vidéos privées hebdomadaires,

b) de les tenir au courant à un rythme hebdomadaire de la progression du cours et de leurs enfants (un courriel et une séquence filmée par cours),

c) de montrer aux familles les progrès faits par leurs enfants et de discuter de la suite à donner à la fin des cours.

4.3. Comment mesurer l'efficacité d'une méthode ?

Une expérience permettant de dire qu'une méthode est efficace ou meilleure qu'une autre méthode nécessite un contrôle des paramètres tels que la motivation, les langues maternelles de chaque apprenant, la participation des parents, le nombre d'heures d'exposition à la langue, l'âge, les révisions à la maison, etc. Ces mêmes paramètres doivent être partagés à l'identique par deux groupes appariés (cible et contrôle). C'est très difficile à réaliser dans les salles de classe actuellement, même si des chercheurs s'y essaient avec bravoure (Krzonowski, Ferragne & Pellegrino, 2016). De plus, il faut que les méthodes à comparer partagent exactement le même but, qui est ici d'acquérir peu de mots, mais avec une bonne prononciation.

4.4. Projets

Différents projets sont en cours :

– le dépôt d'un **projet ANR Jeune chercheur** avec deux buts : 1) établissement d'une base de données audio avec des enfants anglophones d'un âge comparable, permettant d'ajouter des voix d'enfants comme références et d'ouvrir la voix à des mesures de distance automatique et 2) expériences dans deux écoles permettant d'évaluer par des tests hebdomadaires la progression des enfants les acquis à court, moyen et à long terme, pour régler le nombre de répétitions

nécessaires. Notre expérimentation sera conçue sur le modèle de celle de Melnik & Peperkamp (2021).

– les différences entre phonèmes de l'anglais et du français proches mais non identiques seront mises en figurines, comme la différence entre la monophtongue /i/ française et le /iː/ et /ɪ/ anglais, ainsi que des gestes illustrant les différentes habitudes articulatoires, comme une diminution, voire suppression de l'arrondissement anticipatoire des lèvres dans le contraste anglais « peel-pool » par rapport au contraste français « pile-poule », etc.

– la création de comptines sur une base phonétique. Nous avons entamé une réflexion sur le recours à des comptines originales et mises en musique, qui seraient fondées sur un ou deux schémas accentuels répétitifs pour chaque leçon. Par exemple, pour la leçon n° 2 : « Let's play a game » est scandée OOoO (où O est accentuée et o inaccentuée)[35].

– un projet de conte philosophique, illustré et sonorisé, est en cours d'écriture par Maëlle Amand et Éloïse Amand. Il s'intitule *The Purple Bird Travels the World* et relate les aventures d'un petit oiseau mauve qui voyage et fait des rencontres. Les prénoms des personnages comportent la même voyelle que l'animal désigné : Dolly the orange fox, Myrtle the purple bird, Tucker the funny duck, etc.

Conclusion

La méthode que nous proposons pour obtenir la meilleure prononciation possible dès le début de l'apprentissage est préalable et complémentaire des méthodes communicatives et interactionnelles, et ne devrait pas leur être opposée. Une bonne prononciation facilite la communication et désinhibe les apprenants[36].

---

[35] Bouvet (2021, p. 113) suggère le recours aux comptines traditionnelles anglo-saxonnes comme « Humpty Dumpty » ou « Twinkle Twinkle », auxquelles il suggère d'associer une gestuelle particulière.

[36] Nous remercions le LabEx EFL pour le soutien au projet et les relecteurs anonymes de cet article pour leurs remarques et suggestions très utiles. Notre gratitude va également à Maëlle Amand pour sa participation et sa créativité, à Sachie Shioya pour les illustrations et à Gabriel Colomer pour l'écriture du script Python et la musique de fond utilisées pour les vidéos dédiées à l'apprentissage passif par itération.

Références

ADAMCZEWSKI Henri & KEEN Denis, 1973, *Phonétique et phonologie de l'anglais contemporain*, Paris, Armand Colin.

ALAZARD-GUIU Charlotte, ROA Leonardo Contreras, FERRANÉ Isabelle, FOY Marie Philippart de, FONTAN Lionel, YASSINE-DIAB Nadia & PINQUIER Julien, 2022, « Apport du geste dans l'acquisition de la prononciation en L2 : quand la réalité ne correspond pas aux attentes », 34èmes Journées d'Études sur la Parole (JEP 2022), disponible sur : https://hal.science/hal-03855455 (consulté le 3 mai 2023).

ARLEO Andy, 1993, "Should IUT students learn the International Phonetic Alphabet?", *Cahiers de l'APLIUT*, vol. 12, n° 4, p. 39–52.

BANCROFT W. Jane, 2005, *Suggestopedia and Language*, Routledge.

BEUZON Sylvie, BOUCÉ Stéphane, GARCIA Emilie, KESKPAIK Saskia & MARCHOIS Corinne, 2013, *L'évolution des compétences en anglais, en espagnol et en allemand des élèves en fin de collège*, Ministère de l'Éducation Nationale, Secrétariat général, Direction de l'évaluation, de la Prospective et de la Performance.

BOERS Frank & LINDSTROMBERG Seth, 2005, "Finding ways to make phrase-learning feasible: The mnemonic effect of alliteration", *System* [en ligne], vol. 33, n° 2, p. 225-238.

BOTTINEAU Didier, 2022, « Vers une phonologie incarnée et distribuée, motrice, mémétique et énactive », *Signifiances (Signifying)* [en ligne], vol. 6, n° 1, p. 1-39 (consulté le 23 avril 2023).

BOUTIN Gérald, 2004, « L'approche par compétences en éducation : un amalgame paradigmatique », *Connexions* [en ligne] vol. 81, no. 1, p. 25-41. (consulté le 28 novembre 2022).

BOUVET Pascal, 2021, *La phonologie de l'anglais : à l'école, au collège, au lycée*, Belin Éducation.

CAPLIEZ Marc, 2011, « Typologie des erreurs de production d'anglais des francophones : segments vs. suprasegments », *Recherche et pratiques pédagogiques en langues de spécialité*, *Cahiers de l'Apliut* [en ligne], vol. XXX, n° 3, p. 44-60, disponible sur : http://apliut.revues.org/1645 (consulté le 10 juin 2014).

CAPLIEZ Marc, 2015, « Aperçu de l'enseignement de l'anglais oral en France : statut, perspectives et étude de cas », *Recherches en didactiques*, vol. 19, n° 1, p. 109-128.

CARPENTER Angela C., 2015, "Phonetic Training Significantly Mitigates the Stress 'Deafness' of French Speakers", *International Journal of Linguistics*, vol. 7, n° 3, p. 94.

CAUNA Anaïs, TELLIER Marion & COLÉ Pascale, 2021, « La place des gestes pédagogiques en grande section de maternelle », *in* Journée Scientifique de SFERE-Provence [en ligne], disponible sur https://hal-amu.archives-ouvertes.fr/hal-03392061 (consulté le 3 mai 2023).

CHAN Marsha J., 2018, "Embodied Pronunciation Learning: Research and Practice", *CATESOL Journal* [en ligne], vol. 30, n° 1, p. 47-68. (consulté le 18 avril 2023).

CHEN-GÉRÉ Marie-France, DUMEZ Sylvie, HUART Ruth, TANG Maria & VILLAIN Catherine, 2000, *Anglais 6ème Surf in English*, Bordas.

CNESCO, 2019, « Langues vivantes étrangères : comment l'école peut-elle mieux accompagner les élèves ? », Dossier de synthèse [en ligne], disponible sur : https://www.cnesco.fr/wp-content/uploads/2019/04/190410_Dossier_synthese_Langues_.pdf

DELATTRE Pierre, 1953, « Les modes phonétiques du français », *The French Review*, vol. 27, n° 1, p. 59-63.

DELATTRE Pierre, 1965, *Comparing the phonetic features of English, French, German and Spanish. An interim report*, Heidelberg, J. Groos Verlag.

DESGAGNÉ Serge, 1997, « Le concept de recherche collaborative : l'idée d'un rapprochement entre chercheurs universitaires et praticiens enseignants », *Revue des sciences de l'éducation,* vol. 23, n° 2, p. 371–393.

DIEDRICH Jennifer L., 2010, *Motivating Students Using Positive Reinforcement*, Mémoire de Master [en ligne], disponible sur : https://core.ac.uk/download/pdf/233569351.pdf (consulté le 28 novembre 2022).

DUPOUX Emmanuel, PALLIER Christophe, SEBASTIAN Nuria & MEHLER Jacques, 1997, "A Destressing 'Deafness' in French?", *Journal of Memory and Language*, vol. 36, n° 3, p. 406-421.

EEF, 2018, "Phonics, Education Endowment Foundation–Evidence Summaries", [en ligne], disponible sur : https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/phonics/ (consulté le 2 décembre 2019).

EXARE Christelle, 2017, *Les aspirations intrusives dans l'anglais des apprenants francophones*, Thèse de doctorat, Université Paris 3 Sorbonne Nouvelle.

EXARE Christelle, 2022, "Awareness of glottal settings for the production of /h/-initial and vowel-initial words in French learners of L2 English", *Anglophonia. French Journal of English Linguistics*, n° 33.

FAURE Georges, 1975, *Manuel pratique d'anglais parlé*, Paris, Hachette.

FLEGE James Emil, 1984, "The detection of French accent by American listeners", *The Journal of the Acoustical Society of America*, vol. 76, n° 3, p. 692-707.

FLEGE James Emil, 1987, "The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification", *Journal of Phonetics*, vol. 15, p. 47-65.

FROST Dan, 2022, "Doing Pronunciation Online: An Embodied and Cognitive Approach, Which Puts Prosody First", *RANAM*, vol. 55, p. 11-28.

FROST Dan & HENDERSON Alice, 2013, « Résultats du sondage EPTiES (English Pronunciation Teaching in Europe Survey) : l'enseignement de la prononciation dans plusieurs pays européens vu par les enseignants », *Recherche et pratiques pédagogiques en langues de spécialité*, *Cahiers de l'Apliut*, vol. XXXII, n° 1, p. 92-113.

GAONAC'H Daniel & MACAIRE Dominique, 2019, « Les langues à l'école dès le plus jeune âge », [en ligne] CNESCO, disponible sur : https://www.cnesco.fr/wp-content/uploads/2019/04/190409_Gaonach-Macaire.pdf (consulté le 8 mai 2023).

GATTEGNO Caleb, [1963] 1972, *Teaching Foreign Languages in Schools: The Silent Way,* Educational Solutions Worldwide Inc.

HAGHIGHI Mina & RAHIMY Ramin, 2017, "The Effect of L2 Minimal Pairs Practice on Iranian Intermediate EFL Learners' Pronunciation Accuracy", *International Journal of Research in English Education*, vol. 2, n° 1, p. 42-48.

HALIMI Suzy, 2012, « Apprendre les langues, Apprendre le monde », *Rapport du Comité stratégique des langues présidé par Suzy Halimi.*

HALLÉ Pierre A., BEST Catherine T. & LEVITT Andrea, 1999, "Phonetic vs. phonological influences on French listeners' perception of American English approximants", *Journal of Phonetics*, vol. 27, p. 281-306.

HENDRICKSON James M., 1978, "Error Correction in Foreign Language Teaching: Recent Theory, Research, and Practice", *The Modern Language Journal*, vol. 62, n° 8, p. 387-398.

HERRY-BÉNIT Nadine, 2008, « Apprendre à parler l'anglais à l'école primaire : récit d'une expérience », *TIPA. Travaux interdisciplinaires sur la parole et le langage*, n° 27, p. 75-91.

HERRY-BENIT Nadine & CADET Lucile, 2014, « Fonctionnalité des gestes pour l'enseignement de la phonoprosodie », *in* Marion Tellier et Lucile Cadet (dir.), *Le corps et la voix de l'enseignant : théorie et pratique*, Éditions Maison des Langues.

HERRY-BENIT Nadine & PILLOT-LOISEAU Claire, 2012, « Enseignement et évaluation de la prononciation de l'anglais et du français langues étrangères », *Langages, cultures, sociétés : interrogations didactiques : Colloque Diltec*, Paris, France, p. 60-61.

HILTON Heather, 2022, *Profession enseignant – Enseigner les langues avec l'apport des sciences cognitives*, Hachette Éducation.

HONIKMAN Beatrice, 1964, "Articulatory Settings", *in* David Abercrombie, Dennis Butler Fry, Peter A.D. MacCarthy, & Norman Carson Scott (dir.), *In Honour of Daniel Jones*, Longman, London, p. 73–84.

HORGUES Céline, 2010, *Prosodie de l'accent français en anglais et perception par des auditeurs anglophones*, Thèse de doctorat, Université Paris Diderot-Paris 7.

HOSKOVEC J., 1966, "Hypnopedia in the Soviet Union: a critical review of recent major experiments", *The International Journal of Clinical and Experimental Hypnosis*, vol. 14, n° 4, p. 308-315.

HUART Ruth, 2002, *La grammaire orale de l'anglais*, Gap, Ophrys.

IVERSON Paul, PINET Mélanie & EVANS Bronwen G., 2012, "Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels", *Applied Psycholinguistics*, vol. 33, n° 1, p. 145-160.

JEWITT Carey, 2012, "Multimodal Teaching and Learning", *The Encyclopedia of Applied Linguistics*, John Wiley & Sons, Ltd.

JONES Daniel, 2011, *Cambridge English Pronouncing Dictionary*, Cambridge, Cambridge University Press.

KAPSI Sevasti, KATSANTONI Spyridoula & DRIGAS Athanasios, 2020, "The Role of Sleep and Impact on Brain and Learning", *International Journal of Recent Contributions from Engineering, Science & IT (iJES)*, vol. 8, n° 3, p. 59-68.

KAPSI Sevasti, KATSANTONIS Spyridoula & DRIGAS Athanasios, 2021, "Hypnopaedia or sleep learning: Overview on methods and results", *Research, Society and Development*, vol. 10, n° 7.

KRZONOWSKI Jennifer, FERRAGNE Emmanuel & PELLEGRINO François, 2016, « Perception et production de voyelles de l'anglais par des apprenants francophones : effet d'entraînements en perception et en production », *Actes de la conférence conjointe JEP-TALN-RECITAL 2016*, vol. 1, p. 491-499.

LEGENDRE Jacques, 1995, « Suite d'une mission d'information sur l'enseignement des langues vivantes dans l'enseignement scolaire », Online. Rapport du Sénat - Commission des Affaires Culturelles.

LEGENDRE Jacques, 2003, « L'enseignement des langues étrangères en France », Rapport du Sénat.

« L'état de l'École : 32 indicateurs sur le système éducatif français », [en ligne], disponible sur : http://www.education.gouv.fr/cid57102/l-etat-de-l-ecole-32-indicateurs-sur-le-systeme-educatif-francais.html (consulté le 6 novembre 2014).

LEVIS John & CORTES Viviana, 1995, "Minimal Pairs in Spoken Corpora", *ELT Journal*, vol. 49, n° 2, p. 13.

LILLY Richard & VIEL Michel, 2001, « Représentation et réalité en phonologie anglaise », *Études anglaises*, vol. 54, n° 4, p. 387-400.

LINDSEY Geoff, 2019, *English After RP: Standard British Pronunciation Today*, Springer.

LITTLEWOOD William & YU Baohua, 2011, "First language and target language in the foreign language classroom", *Language Teaching*, vol. 44, n° 1, p. 64-77.

Loi n° 2013-595 du 8 juillet 2013 d'orientation et de programmation pour la refondation de l'École de la République, 2013, Ministère de l'Éducation Nationale et de la Jeunesse [en ligne], disponible sur https://www.education.gouv.fr/loi-ndeg2013-595-du-8-juillet-2013-d-orientation-et-de-programmation-pour-la-refondation-de-l-ecole-5618.

LYSTER Roy, 1998, "Negotiation of Form, Recasts, and Explicit Correction in Relation to Error Types and Learner Repair in Immersion Classrooms", *Language Learning, vol.* 48, n° 2, p. 183-218.

LYSTER Roy & RANTA Leila, 1997, "Corrective feedback and learner uptake", *Studies in Second Language Acquisition*, n° 1, p. 37-66.

MANOÏLOV Pascale, 2019, *Les acquis des élèves en langues vivantes étrangères*, Paris, CNESCO.

MARSLEN-WILSON William D & WELSH Alan, 1978, "Processing interactions and lexical access during word recognition in continuous speech", *Cognitive Psychology*, vol. 10, n° 1, p. 29-63.

MEIRIEU Philippe & HAMELINE Daniel (dir.), 2000, *L'école et les parents : la grande explication*, Paris, Plon.

MELNIK Gerda Ana & PEPERKAMP Sharon, 2021, "High-Variability Phonetic Training enhances second language lexical processing: evidence from online training of French learners of English", *Bilingualism: Language and Cognition*, vol. 24, n° 3, p. 497-506.

Méthodologie du cycle des évaluations disciplinaires réalisées sur échantillon (Cedre) en fin d'école et fin de collège, 2004, Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche.

MICHARI Dylan, 2023, « » Oser dire le nouveau monde » : l'enseignement-apprentissage des accents en classe d'anglais à l'aune du plan langues vivantes », *Journées d'études de l'ALOES – Quels objectifs pour une phonologie au service du sens ?* 31 March 2023.

MOLL Kristina, RAMUS Franck, BARTLING Jürgen, *et al.*, 2014, "Cognitive mechanisms underlying reading and spelling development in five European orthographies", *Learning and Instruction*, vol. 29, p. 65-77.

MOMPEAN Jose A. & FOUZ-GONZÁLEZ Jonás, 2021, "Phonetic Symbols in Contemporary Pronunciation Instruction", *RELC Journal: A Journal of Language Teaching and Research*, vol. 52, n° 1, p. 155-168.

NAGLE Charles L. & BAESE-BERK Melissa M., 2022, "Advancing the state of the art in L2 speech perception-production research: revisiting theoretical assumptions and methodological practices", *Studies in Second Language Acquisition*, vol. 44, n° 2, p. 580-605.

NARCY-COMBES Jean-Paul, TARDIEU Claire, LE BIHAN Jean-Claude *et al.*, 2008, « L'anglais à l'école élémentaire », *Langues Modernes*, vol. 102, n° 4, p. 72-82.

NATTINGER James R. & DECARRICO Jeanette S., 1992, *Lexical Phrases and Language Teaching*, Oxford, Oxford University Press.

ODENDAAL M. S., 1987, "Raising achievement levels by means of Suggestopedia", *Per Linguam*, vol. 3, n° 2.

PAWLEY Andrew & SYDER Frances Hodgetts, 1983, "Two puzzles for linguistic theory: nativelike selection and nativelike fluency", *Language and Communication*, Routledge.

PINSKER Melvin, 1971, "Learning While Relaxing or Sleeping", *Education*, vol. 92, n° 2, p. 51-53

POTAPUSHKINA-DELFOSSE Marie, 2014, *Gestes, contes, théâtre : une approche multimodale de l'anglais pour des élèves débutants à l'école primaire*, Thèse de doctorat, Le Mans.

POURHOSEIN GILAKJANI Abbas & SABOURI Narjes, 2016, "Why Is English Pronunciation Ignored by EFL Teachers in Their Classes?", *International Journal of English Linguistics,* vol. 6, p. 195.

« Première Enquête européenne sur les compétences linguistiques », 2012, [en ligne], disponible sur : http://ec.europa.eu/languages/library/studies/executive-summary-eslc_fr.pdf

RÉVAUGER Guilène, 2021, « L'enseignement de la civilisation du monde anglophone en cours

d'anglais dans le second degré en Polynésie française : pratiques enseignantes et tensions »,
Contextes et didactiques, Revue semestrielle en sciences de l'éducation, n° 18.

WHEATLEY Rikki K., WEST Richard P., CHARLTON Cade T. *et al.,* "Improving behavior
through differential reinforcement: A praise note system for elementary school students", *Education
and treatment of children*, 2009, vol. 32, n° 4, p. 551-571.

ROGERS James, WEBB Stuart & NAKATA Tatsuya, 2015, "Do the cognacy characteristics of
loanwords make them more easily learned than noncognates?", *Language Teaching Research*, vol. 19,
n° 1, p. 9-27.

ROSENSHINE Barak V., 1986, "Synthesis of Research on Explicit Teaching", *Educational
Leadership*, vol. 43, n° 7, p. 60-69.

SCHMITT Norbert, 2004, *Formulaic Sequences: Acquisition, Processing and Use*, John Benjamins
Publishing.

SCHWIEREN Juliane, BARENBERG Jonathan & DUTKE Stephan, 2017, "The Testing Effect
in the Psychology Classroom: A Meta-Analytic Perspective", *Psychology Learning & Teaching, vol.* 16,
n° 2, p. 179-196.

« Socle commun de connaissances, de compétences et de culture », 2015, Ministère de l'Éducation
Nationale et de la Jeunesse, [en ligne], disponible sur :
https://www.education.gouv.fr/bo/15/Hebdo17/MENE1506516D.htm

TAYLOR Alex & MANES-BONNISSEAU Chantal, 2018, « Propositions pour une meilleure
maîtrise des langues vivantes étrangères », Ministère de l'Éducation nationale et de la Jeunesse,
disponible sur : https://www.education.gouv.fr/cid133908/propositions-pour-une-meilleure-
maitrise-des-langues-vivantes-etrangeres.html

TELLIER Marion, 2010, « Faire un geste pour l'apprentissage : Le geste pédagogique dans
l'enseignement précoce », *in* C. Corblin & J. Sauvage (dir.), *L'apprentissage et l'enseignement des langues
vivantes à l'école. Impacts sur le développement de la langue maternelle*, Paris, L'Harmattan, collection
« Enfance & Langages », p. 31-54.

THOMSON Ron I., 2018, "High Variability [Pronunciation] Training (HVPT) : A proven
technique about which every language teacher and learner ought to know", *Journal of Second Language
Pronunciation*, vol. 4, n° 2, p. 208-231.

TORTEL Anne, 2009, *Evaluation qualitative de la prosodie d'apprenants français : apport de paramétrisation
prosodiques*, Thèse de doctorat, Université de Provence-Aix-Marseille I.

TUAN Luu Trong, 2010, "Teaching English Discrete Sounds through Minimal Pairs", *Journal of Language Teaching and Research*, vol. 1, n° 5, p. 540-561.

« Un plan d'actions pour une meilleure maîtrise des langues vivantes étrangères », 2022, Éduscol – Ministère de l'Éducation nationale et de la Jeunesse – Direction générale de l'enseignement scolaire.

VOISE Anne-Marie, 2010, « Enseigner la phonologie de l'anglais aux futurs professeurs du primaire », *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Apliut*, vol. XXIX, n.° 2, p. 11-24.

VOISE Anne-Marie, 2018, « Corps – accords : une approche holistique et transdisciplinaire des langues à l'école maternelle », *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Apliut*, vol. 37, n° 2.

VULETIĆ Branko, 1965, « Système des fautes et correction phonétique des français qui apprennent l'anglais », *Revue de Phonétique Appliquée*, 1965, n° 1, p. 1-39.

WELLS John, 2008, *Longman Pronunciation Dictionary*, Pearson Longman.
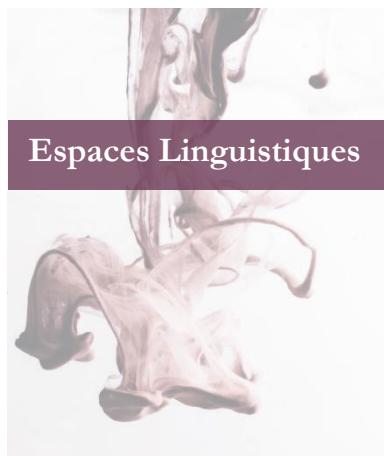
Realisation, elision and intrusion of /h/ in French learners of English

Réalisation, élision et intrusion de /h/ chez les apprenants francophones en anglais

Marc CAPLIEZ

Université du Littoral Côte d'Opale, UR 4030

marc.capliez@univ-littoral.fr

Résumé : Comme les systèmes phonologiques de l'anglais et du français se distinguent à bien des égards, à la fois au niveau segmental et au niveau prosodique, les apprenants francophones rencontrent d'importantes difficultés avec la prononciation anglaise comme, par exemple, la fricative glottale /h/, absente de l'inventaire phonémique du français – à l'exception de l'onomatopée *haha* qui traduit le rire. À la suite des études d'Exare (2017) et de Kamiyama *et al.* (2011), l'objectif de cet article est d'apporter des preuves empiriques de la réalisation, l'élision et l'intrusion de /h/ dans des productions d'anglais d'étudiants francophones non-spécialistes de l'anglais. L'analyse auditive du corpus, composé de trois tâches (lecture de mots, lecture de phrases et discours spontané), fait ressortir quatre profils d'apprenants : (1) le /h/ est absent ; (2) le /h/ est correctement réalisé ; (3) le /h/ est correctement réalisé et le /h/ intrusif est récurrent ; (4) seul le /h/ intrusif est produit. L'étude montre une certaine cohérence intra-locuteur, mais il existe en parallèle une variabilité importante selon le type de tâche, ainsi qu'une différence entre les femmes et les hommes. Ces résultats mettent en exergue l'intérêt d'adapter les stratégies d'apprentissage à ces différents profils.

Mots clés : acquisition L2, anglais, français, fricative glottale, consonne

Abstract: As the phonological systems of English and French diverge both at the segmental level and at the prosodic level, French learners encounter difficulties with English pronunciation. One example is the glottal fricative /h/ since it is absent from the French phonemic inventory–with the exception of the onomatopoeia "*haha*" to express laughter. Following studies by Exare (2017) and Kamiyama *et al.* (2011), this paper aims to provide further empirical evidence of the realisation, elision, and intrusion of /h/ in French speakers' English productions. The learners were university students who did not major in English. The participants were given three tasks, i.e., word reading, sentence reading, and spontaneous speech. Auditory analyses were then carried out by the researcher. Results show four speaker-profiles: (1) /h/ is absent altogether; (2) /h/ is accurately realised; (3) /h/ is accurately realised in addition to regular occurrences of intrusive /h/; (4) only intrusive /h/ occurs. Although the speaking task has an influence on intra-speaker variability, intra-speaker consistency has been observed, with participants belonging to one of the above learner-profiles. Also of interest were the differences between men and women. These results suggest that different teaching strategies should be established to better suit the needs of the different profiles.

Keywords: L2 acquisition, English, French, glottal fricative, consonant

Introduction

Over the last few decades, there has been an increasing interest in research on the acquisition of English phonology by non-native speakers. Despite a geographical and historical proximity, the phonological systems of English and French diverge in many ways, both at the segmental level (i.e., vowels and consonants) and at the prosodic level (i.e., stress, rhythm and intonation). French speakers encounter major difficulties with English pronunciation, resulting in stereotypes and mockery such as replacing the English post-alveolar approximant /ɹ/ by the French uvular fricative /ʁ/ in the realisation of the grapheme <r>, and systematically omitting the glottal fricative /h/ as this sound is not part of the French phonemic inventory. Indeed, the consonant <h> is problematic on many levels; not only is it always silent in French, but teachers and researchers have also observed that some French learners add a [h] sound in front of vowel-initial words. However, there is still a need for empirical evidence, along with qualitative and quantitative measurements of these elision and intrusion phenomena.

First, we will present an overview of research on the acquisition of English phonemes by French speakers by comparing the phonological systems of the two languages and introducing theoretical frameworks of L2 (i.e., second/foreign language) segmental acquisition. We will then focus on the case of the glottal fricative /h/ and the graphic consonant <h>. We have attempted to quantify the realisation of /h/ and compare its systematic omission and intrusion in French speakers' English productions through an auditory analysis of spoken data that were collected from a cohort of French students. Finally, we will discuss our observations, including a comparison of female and male students' performances, drawing conclusions regarding the implications for the teaching of English as a foreign language.

1. The voiceless glottal fricative /h/

1.1. The phonemic inventories of English and French

English and French have diverging phonemic inventories despite sharing the same alphabet. Standard British English has 44 phonemes, including 22 consonants and two semi-consonants, whereas Standard Parisian French has 37, including 18 consonants and three semi-consonants. Not only are their vowel systems different in number and quality, their consonants also present major phonological and phonetic differences (Capliez, 2011). The typical examples that are used to illustrate the differences between the consonants of the two languages are the post-alveolar approximant /ɹ/ (commonly represented by the IPA symbol /r/) in Southern British English

(SBE), or the retroflex approximant /ɻ/ in General American (as in *red*), and the dental fricatives /θ/ and /ð/ (as in *thing* and *this*, respectively), which are absent in the French phonemic inventory–although the latter sounds are phonetically produced by speakers with a lisp (i.e., *cette* "this" will be pronounced [θɛt] instead of /sɛt/). Likewise, the sound /h/ is not a contrastive sound and therefore, not part of the French phonemic inventory.

The graphophonemic rules of English and French have little in common. When the grapheme <h> combines with another grapheme, the resulting sounds often differ. For example, the complex grapheme <ch> is pronounced /tʃ/ in English (e.g., *check* /tʃek/) and /ʃ/ in French (e.g., *chose* "thing" /ʃoz/), and the grapheme <th> corresponds to the dental fricatives /θ/ and /ð/ in English, but to the voiceless plosive /t/ in French. Silent consonants also occur in different contexts; thus, <l> is automatically silent in the sequences <alk>, <alf> and <alm> in English (e.g., *walk*, *half*, *calm*), whereas it is rarely silent in French. As far as the simple grapheme <h> is concerned, it is present in both languages. However, it does not behave in a similar way. While it is realised as a voiceless glottal fricative /h/ in English, it is always silent in French.

1.2. The case of the grapheme <h>

The letter <h> is silent in French, excluding breathlessness and onomatopoeias expressing such emotions as exasperation or relief where a [h] sound can be heard (e.g., *ha!*). However, a distinction is traditionally made between unaspirated "h" and aspirated "h". Although the latter has no phonetic reality nowadays, contrary to English so-called "aspirated" plosive consonants, it represents a hiatus involving the final vowel of the preceding word and the initial vowel of the "h" word (see Encrevé, 1988 for more details). Accordingly, the contraction and liaison phenomena do not occur with a word beginning with an aspirated "h", whereas they do with unaspirated "h". For example, the masculine noun *homme* "man" /ɔm/ begins with an unaspirated "h", and thus the definite article le will become *l'* (*l'homme* "the man") as it always does before a vowel (cf. Encrevé, 1988). Similarly, the indefinite article *un*, normally pronounced as the nasal vowel / œ̃ / (or increasingly /ɛ̃/), will have an audible /n/ representing the liaison and avoiding the hiatus (*un homme* "a man" / œ̃ n ɔm/), French usually preferring a CV (i.e., onset consonant + vowel) syllable structure (Adda-Decker *et al*, 2002). On the contrary, the word *héros* has an aspirated "h"; in that case, the definite article will remain *le* and no consonant is added at the end of the indefinite article (i.e., *le héros* "the hero" /lə eʁo/, *un héros* "a hero" / œ̃ eʁo/). As in any other words beginning with

a vowel sound, a glottal stop may be added in order to give it more emphasis. Some dictionaries indicate aspiration with the diacritic ˈ, even though it is not related to lexical stress.

In SBE, <h> is seldom silent. Most of the time, it corresponds to the voiceless glottal fricative /h/, as explained above. Only a few exceptions are to be found, in the onset of the words *hour*, *honest*, *hono(u)r*, *heir* (and their derived words: *hourly*, *honesty*, *honourable*, *heirloom*, etc.) and in medial position in some words such as *vehicle*. In those words, the <h> is silent, thus requiring the indefinite article *an* instead of *a*, and the weak form of the definite article *the* is pronounced /ði/ instead of /ðə/. To that list can be added the word *herb* in General American, some words beginning with an unstressed syllable (e.g., sometimes *an his ˈtorian*), and direct French loanwords such as *haute couture* and *hors d'oeuvre*. Excluding the natural reduction of unstressed grammatical words such as *her* and *have*, there are a few English accents where "h-dropping" occurs, such as Scouse English, Cockney English and Welsh English (see Wells, 1970), but this phenomenon is usually stigmatised (Roach, 2009).

1.3. The acquisition of /h/ by French speakers

The acquisition of the phonemes of a foreign or second language (L2) has been widely studied. Flege's Speech Learning Model (SLM) has been developed to account for the influence of the mother tongue (L1) in the acquisition process of the L2. According to this model (Flege, 1995), L2 phonemes are easier to acquire for a speaker whose L1 phonemic inventory does not have any similar phonemes, whereas L2 phonemes which are closer to L1 phonemes are more difficult to acquire. Hence, the acquisition of /h/ should be easier to acquire for French learners because it is not part of the French inventory. However, according to Eckman's Markedness Differential Hypothesis (MDH, Eckman, 1977; Eckman, 2008), the "markedness" of a linguistic phenomenon, that is to say the degree of rarity of its occurrence within or between languages, is a factor of difficulty in L2 acquisition, and therefore an unmarked sound (that is, a sound that can be found in most languages) is easier to produce for a non-native speaker than a marked sound (i.e., more specific to one language). Based on this model, /h/ should be difficult to acquire for French learners of English.

Considering these models of L2 phonology acquisition, one may predict that the acquisition of the phoneme /h/ is bound to be problematic for French learners (see Janda & Auger, 1992). De Launay (1993) presented some French learners with pairs of English words such as *hair*/*air* and concludes: "students usually find it both illuminating and fun to discover these words" (*Ibid.*, p. 134). Brown (1988, 1991) developed the idea of functional load of the members of a minimal

pair, based on the cumulative frequency of the two members. Considering that pairs of words such as *hair/air* or *hate/ate* bear relatively high functional load, their perception and production by non-native speakers will be all the more difficult. Although some pairs are unlikely to cause misperception or misunderstanding because of their different uses or natures (e.g., the noun *hat* vs. the preposition *at*), others may lead to misinterpretations which may hinder communication (e.g., *ate* vs. *hate*; Capliez and Guendouz, 2021).

The absence of phonemic /h/ in French is bound to result in its elision in French speakers' English productions, particularly if one considers the influence of the spelling and graphophonemic rules of the L1. Kamiyama, Kühnert and Vaissière (2011) asked some French learners of English to read aloud the nursery rhyme *Humpty Dumpty* and found that out of 37 students, 14 of them dropped /h/ at least once in the word "Humpty". More surprisingly, a widespread phenomenon among French learners of English consists in inserting /h/ between a word ending in a vowel and a word beginning with a vowel (*hiatus*); that is known as "intrusive /h/", or "h-epenthesis" (De Launay, 1993; Exare, 2017; Janda & Auger, 1992). As explained above, the French language normally prefers a CV(C) syllable structure (Adda-Decker *et al*, 2002) and often resorts to linking phenomena in hiatuses, such as the pronunciation of the <n> in a final nasal vowel before another vowel (e.g., *un homme* "a man" / œ̃ n ɔm/, *mon amour* "my love" /mɔ̃ n amuʁ/; see Encrevé, 1988). Furthermore, liaison can result in the addition of a graphic <t> as in the interrogative form of *il y a* "there is" → *y a-t-il* "is there", or *il va* "he is going" → *va-t-il* "is he going", which avoids the sequence of vowels. Some erroneous instances can thus be heard; for example, some French speakers are tempted to avoid a hiatus by inserting a consonant sound in a phrase such as *les gens qui ont* "the people who have", becoming *les gens qui z'ont*, and even humorously in *moi aussi* "me too", becoming *moi z'aussi*.

As is noted in Capliez and Guendouz (2021), intrusive /h/ is a commonly observed phenomenon among French learners of English, and it may be seen as evidence of a certain amount of effort from the learner who is attempting to pronounce the target language accurately; Janda and Auger (1992) categorise it as being a form of "qualitative hypercorrection". Exare's study (2017) focused on glottalisation and aspiration in French learners' L2 speech and hypothesised that these occurrences of hypercorrection may result from a lack of assimilation of the [ʔ] ~ [h] contrast, phonological repair of vowel-initial environments, or an articulatory process involving the intrusive gesture of glottal opening.

The extent of the intrusive /h/ phenomenon along with the expected elision of phonemic /h/ deserve particular attention in empirical studies involving the acquisition of English phonology by

French speakers. The present study investigates the realisation, elision and intrusion of the sound /h/ in French speakers' L2 English pronunciation based on an auditory analysis.

2. Methodology

2.1. Participants

The recordings of 94 French learners of English as a foreign language were collected and analysed. Among the cohort of French speakers were 42 men and 52 women, aged between 18 and 23. All of them were undergraduate students who had been learning English since secondary school, ranging from eight to ten years of learning in total. They specialised in mathematics, physics, chemistry, or computer science. None of them were English majors or familiar with the study of English phonetics and phonology. Their academic exposure to English at the time of recording was limited to two weekly hours and included listening, speaking, writing, and speaking activities. All the students took a test called ELAO (Efficient Language Assessment Online), revealing that their levels of English proficiency were situated between B1 and B2, according to the Common European Framework of Reference for Languages.

2.2. Data collection

As is suggested by John and Cardoso (2009), the extent of the /h/ problem may differ depending on the type of speech. Thus, every student of our study was asked to take part in three types of recording sessions, presented as part of their English course: word reading, sentence reading and spontaneous speech. They were not informed of the objectives of the experiment.

For the first task, they had to read aloud 30 isolated words that randomly appeared one by one on a computer screen. The words included 12 words beginning with vowels (e.g., *eat, actor, either, old*, etc.), 12 words beginning with a phonemic /h/ (e.g., *hand, home, hut*, etc.) and 6 distractors, that is, words beginning with other consonants than <h> (e.g., *dog, leave, book*, etc.). The former two categories included such pairs as *ill* and *hill*. No exceptional word where <h> is silent (e.g., *hour, heir*) was included. For the sentence-reading task, the students had to read aloud 10 sentences that randomly appeared one by one on a computer screen. In addition to a few distractors and grammatical words beginning with vowels (e.g., *My cousin now lives in London*), some sentences contained target lexical words beginning with vowels (e.g., *The author of the book is under arrest*), some contained target words beginning with /h/ (e.g., *A horrible thing happened to Mary's husband*), and other sentences included both (e.g., *His name is Harry Evison*). Special care was taken to ensure that

no sentence resulted in a tongue twister that even a native or experienced speaker might have difficulty in uttering. The full list of stimuli for both reading tasks is provided in the appendix.

In order to collect spontaneous speech from the students, they were asked to describe their "ideal trip" for approximately two minutes. They were free to speak about a planned trip, an imaginary journey, or even a past travelling experience. They had a few minutes to think about it and take some notes on a piece of paper, but not entire sentences. Despite the presence of the author to give the students instructions and start the recording device, they were not interrupted while they spoke.

In total, 3,290 files were collected and used for auditory analysis, which corresponded to the following data (distractors excluded) and targets:

- 24 words (times 94 students = 2,256 sound files).

Targets: initial vowels = 12 (total = 1128); initial <h> = 12 (total = 1128).

- 10 sentences (times 94 students = 940 sound files).

Targets: initial vowels = 20 (total = 1,880); initial <h> = 16 (total = 1,504).

- 94 spontaneous speech samples.

2.3. Research questions

Based on Flege's SLM, one may hypothesise that the acquisition of the English phoneme /h/ by L1 French speakers should be facilitated by its absence in their native language. However, grapho-phonemic rules regarding <h> in French along with the avoidance of a hiatus in French should be factored in since cross-linguistic transfers are likely to occur. How far do these linguistic transfers take place and what strategies do learners use to avoid hiatus?

3. Results

In this section, we will present the results and statistics[37] of the auditory analyses conducted on the sound files that we collected, first for word reading, second for sentence reading, and third for spontaneous speech. We will also look at the number of speakers concerned by the different phenomena (i.e., the accurate realisation of the targets, the dropping of /h/, and the intrusive /h/ phenomenon).

---

37 Many thanks to Maelle Amand, who significantly helped us in the statistical analyses of our results.

3.1. Word reading

Out of the 2,256 sound files that were collected from the word-reading task, 1,128 contained target
initial vowels and 1,128 target <h> graphemes (i.e., for each type, 12 targets times 94 students).
The following graph shows the realisations of the target initial vowels and target /h/.

Figure 1. Realisations of initial vowels and initial /h/ (word reading)
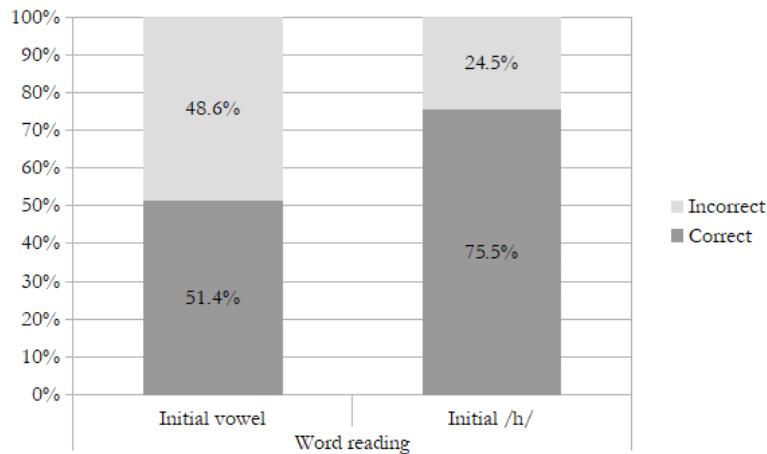
The auditory analyses revealed that 580 target vowels were produced correctly (i.e., 51.4%; 288 by
female speakers and 292 by male speakers), that is to say the initial vowels were respected, whereas
548 words contained an initial intrusive /h/ instead (i.e., 48.6%; 336 by female speakers and 212 by
male speakers). A proportion test allowed us to conclude that the difference between the two types
of realisation was not significant, $z = 1.34$, $p = 0.17$.

Out of 1,128 target initial /h/ tokens, 852 (i.e., 75.5%; 480 by female speakers and 372 by male
speakers) were correctly produced, that is to say the target /h/ phoneme was accurately realised.
In 276 words (i.e., 24.5%; 144 by female speakers and 132 by male speakers), however, the target
/h/ was dropped. A proportion test allowed us to conclude that the difference between the two
types of realisation was significant, $z = 24.25$, $p < 0.00001$.

Figure 2 below shows the performances of men and women.

Figure 2. Men vs women (word reading)
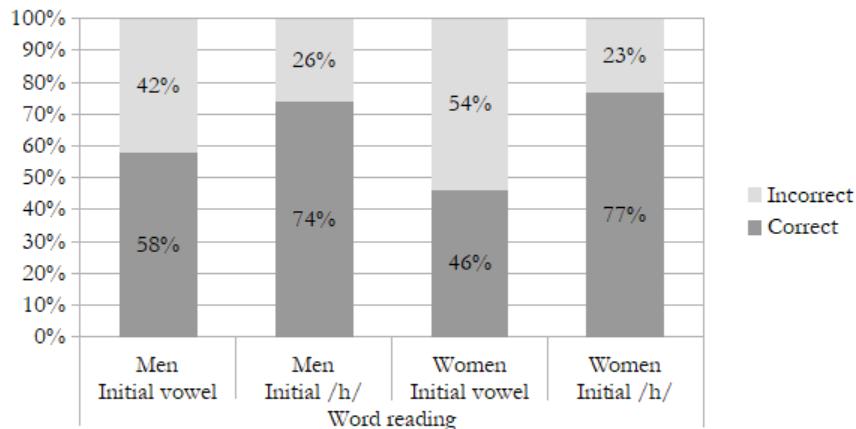
The difference between the male speakers' and the female speakers' realisations of initial vowels is significant: z = -3.93, p = 0.00008. However, the difference is not significant in their realisations of initial /h/: z = 1.21, p = 0.22.

3.2. Sentence reading

The 940 sentences of the second production task contained 3,384 targets, which included 1,880 initial vowels (being the context where the intrusive /h/ phenomenon can be observed and measured) and 1,504 words with an initial <h> /h/.

Figure 3 below shows the realisations of the target initial vowels and target /h/.

Figure 3. Realisations of initial vowels and initial /h/ (sentence reading)
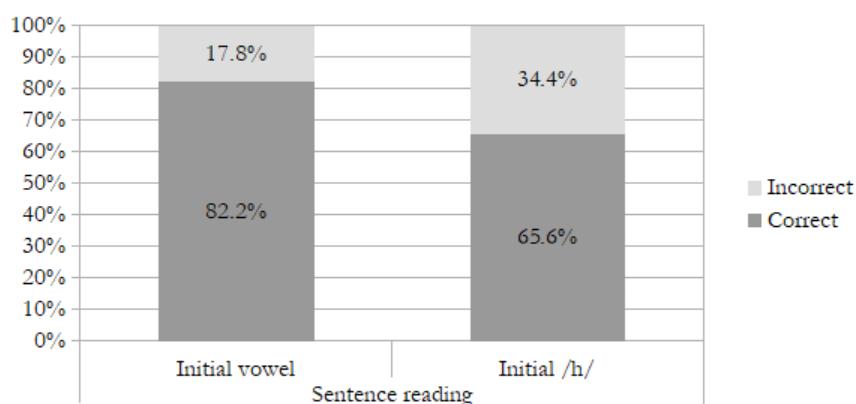
Out of 1,880 targets initial vowels (i.e., 20 targets times 94 speakers), 1,545 (i.e., 82.2%; 821 by female speakers and 724 by male speakers) were correctly realised, whereas 335 (i.e., 17.8%; 219 by female speakers and 116 by male speakers) contained an intrusive /h/. A proportion test allowed

us to conclude that the difference between the two types of realisation was significant, $z = 39.46$, $p < 0.00001$.

Out of 1,504 target initial /h/ tokens (i.e., 16 targets times 94 speakers), 986 (i.e., 65.6%; 685 by female speakers and 301 by male speakers) were correct, which means that the speakers accurately produced /h/ at the beginning of the words; in 518 words (i.e., 34.4%; 147 by female speakers and 371 by male speakers) the /h/ was systematically dropped. A proportion test allowed us to conclude that the difference between the two types of realisation was significant, $z = 17.06$, $p < 0.00001$.

Figure 4 below shows the performances of men and women.

Figure 4. Men vs women (sentence reading)
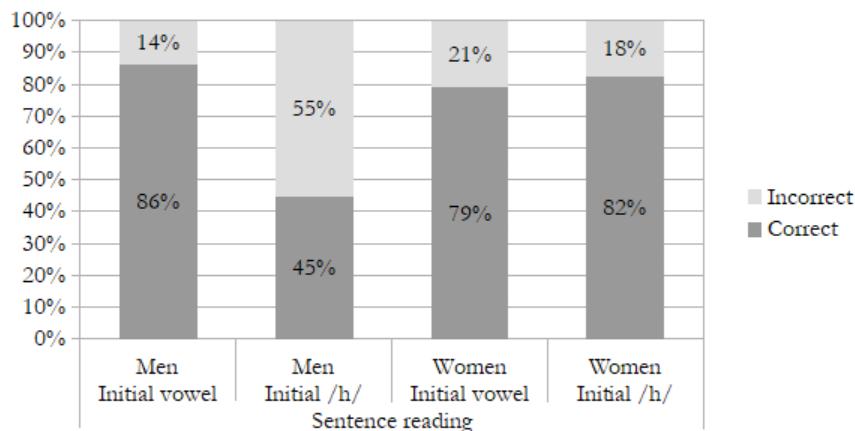
The difference between the male speakers' and the female speakers' realisations of initial vowels is significant: $z = 15.23$, $p < 0.00001$. It is also significant in their realisations of initial /h/: $z = -4.08$, $p < 0.00001$.

3.3. Spontaneous speech

The auditory analyses of the 94 spontaneous speech samples enabled us to observe a considerable amount of intra-speaker consistency. Indeed, four recurring patterns of occurrences of /h/ and realisation of initial vowels emerged in the speakers' L2 productions, resulting in the following typology:

- Pattern 1: the /h/ sound is altogether absent.

- Pattern 2: the initial grapheme <h> is correctly realised as the voiceless glottal fricative /h/.

- Pattern 3: the initial grapheme <h> is correctly realised and intrusive /h/ regularly occurs.

- Pattern 4: only intrusive /h/ occurs; target /h/ (i.e., the realisation of the grapheme <h>) is absent.

In Table 1 below, the numbers of speakers concerned by each of the four patterns are provided. For clarity and comparison purposes, the numbers have also been converted to percentages, and the difference between male and female speakers is given.

Table 1. Distribution of /h/ in the learners' L2 speech

	Absent /h/	Correct /h/	Correct /h/ + intrusive /h/	Intrusive /h/ only	Total
Speakers	35 (37.2%)	29 (30.9%)	20 (21.3%)	10 (10.6%)	94
Male speakers	22 (52.4%)	8 (19%)	6 (14.3%)	6 (14.3%)	42
Female speakers	13 (25%)	21 (40.4%)	14 (26.9%)	4 (7.7%)	52

The following graph illustrates the different patterns that were observed:

Figure 5. Distribution of /h/ in the learners' L2 speech
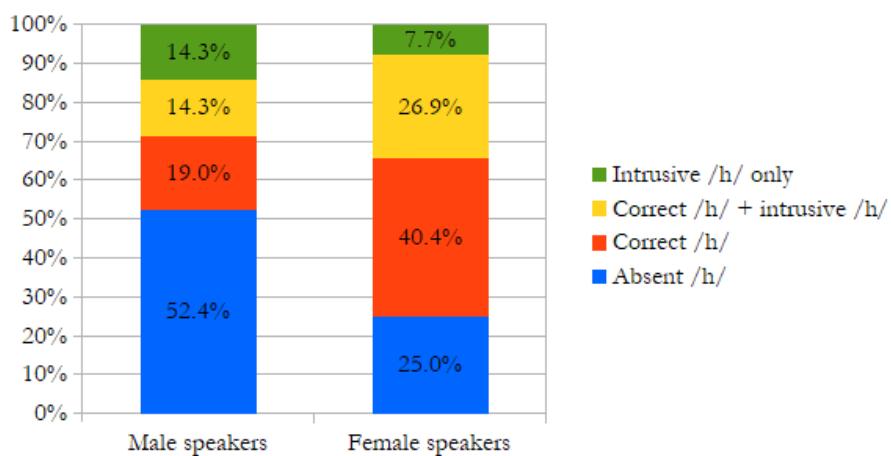
Overall, the prominent pattern of /h/ realisation is Pattern 1, which means that a majority of the speakers (i.e., 37.2%) produced no /h/ at all, followed by 30.9% who realised /h/ correctly (Pattern 2), 21.3% who produced both correct /h/ and intrusive /h/ (Pattern 3), and 10.6% who exhibited intrusive /h/ only, in vowel-initial words (Pattern 4). Various proportion tests allowed us to observe a significant difference between Pattern 1 and Pattern 3 ($z = 2.40$, $p = 0.016$), Pattern 1 and Pattern 4 ($z = 4.27$, $p < 0.0001$), Pattern 2 and Pattern 4 ($z = 3.41$, $p = 0.0006$), and

Pattern 3 and Pattern 4 ($\chi = 1.99$, $p = 0.04$). The differences between Pattern 1 and Pattern 2, and between Pattern 2 and Pattern 3 were not significant (respectively: $\chi = 0.92$, $p = 0.35$ and $\chi = 1.49$, $p = 0.13$).

A chi-square test of independence was performed to examine the relation between gender and the different realisations of /h/. The relation between these variables was significant, $X^2 (3, N = 84) = 10.8$, $p = 0.012$.

3.4. Discussion

According to predictions based on the SLM and the Markedness Hypothesis, the acquisition of the English glottal fricative /h/ should not be problematic for French speakers, whose L1 phonemic inventory does not possess that phoneme, and therefore a new mental category should be created. Nevertheless, the fact that the participants in our study were late learners and had been exposed to English for several years was bound to have an impact on their production. Our auditory analyses revealed notable variability depending on the type of speech (i.e., word reading, sentence reading, spontaneous speech), and the widespread observation of the intrusive /h/ phenomenon by some English teachers and researchers has indeed been verified (De Launay, 1993; Exare, 2017). Our results confirm John and Cardoso's findings (2009), which showed that h-epenthesis was more frequent in more formal styles of speech (i.e., laboratory-controlled reading tasks) among French learners of English. Another significant factor appeared in this study: proportions of correct realisations seemed to depend on the learners' gender.

The auditory analyses of the **word-reading task** did not reveal a significant difference between the correct realisations of initial vowels and the intrusive /h/ occurrences. Furthermore, most speakers seemed to be able to realise the correct target /h/ sounds (75.5%). The fact that this was a word-reading task may explain those results; the learners' attention was focused on precise items which may have helped them realise the targets more accurately.

For the **sentence reading task**, there were significantly fewer occurrences of intrusive /h/ than correct realisations of initial vowels (82.2%). Similarly, a significant number of words beginning with <h> were pronounced with /h/ (65.6%). In other words, there were more correct realisations in both reading tasks than there were errors. The participants' motivation and level of English may explain these results.

The intrusive /h/ phenomenon occurred more in female speakers, and so did the correct realisation of the target /h/, in both reading tasks. This observation may suggest that female speakers are more likely to aim at a more English-sounding pronunciation, /h/ being a "typically

English" sound to French learners. It is impossible to generalise, however, and future studies should further examine the impact of the gender effect, taking into account other factors such as age and level of proficiency.

The analyses of the **spontaneous speech** samples showed a general tendency for the French speakers to drop the /h/ at the beginning of words. Contrary to the previous two controlled tasks involving carefully-selected stimuli, the spontaneity, the lack of preparation and mental correction, and the influence of the L1 seem to have played an important role in the French speakers' productions. Although the intrusive /h/ phenomenon often occurred alongside correct realisations of /h/, its sole occurrence (i.e., what we have named Pattern 4), however unlikely it may have seemed, could still be heard from a noteworthy number of speakers. Such an observation is all the more unexpected as one may be inclined to believe that a French speaker, whose L1 does not include the /h/ phoneme, will either omit it altogether, or realise it–at least partly–accurately in natural, unprepared L2 speech.

As regards the difference in performance between male and female speakers, it was also significant in spontaneous speech, the latter being more likely to exhibit an intrusive /h/ as well as having a correct realisation of /h/. Once again, we believe that the amount of personal effort to make L2 pronunciation sound more accurate may explain these results. By contrast, many male speakers exhibited systematic dropping of /h/. The feeling of embarrassment and the possible difference in proficiency level might explain this gender difference. It is worth noting, however, that a considerable number of male speakers who did not produce /h/ systematically added an audible glottal stop [ʔ] before vowel-initial words in all three tasks; we did not observe that in female speakers. In some participants' productions, the phenomenon even occurred in vowel-initial grammatical words (e.g., *he is* pronounced [ʔi ʔiz] or [hi hiz]).

Drawing a parallel between intrusive /h/ in female speakers and glottal stops in male speakers, both seem to result from a special effort on the part of the learners to reach a native-like pronunciation. Vaissière (1986, p. 537) indeed points out that aspiration and glottalisation "have a common characteristic: a greater tensing of at least one of the articulators, the vocal folds [...]". The frequency of intrusive /h/ found in the data suggests that it is not an articulatory difficulty, despite its absence from the French phonemic inventory. However, using the sound in the right place in English is a challenge.

One of the first steps for the EFL teacher is to simply make learners aware of the existence of the phoneme as early as possible, including by giving minimal pairs–words and phrases–as examples. Certain exercises may help learners improve their pronunciation at whatever their level. Practising

Espaces Linguistiques N°5 | 2023 : Exploring norms and variation in L1/L2 pronunciation
https://www.unilim.fr/espaces-linguistiques - ISSN : 2729-3548

linking may prove to be effective to avoid intrusive /h/ or glottal stops in vowel-initial words and make speech more natural (Capliez & Guendouz, 2021). Such exercises involve the pronunciation of two words as one by inserting a semi-vowel that will avoid /h/ or /ʔ/ epenthesis. For example, the teacher can suggest inserting /j/ in the phrase *my eyes* and using the final consonant /z/ as the first one for the following word, resulting in *my eyes are blue* /maɪ jaɪzə bluː/. The consonant /h/ should therefore be regarded as a teaching priority (Collins & Mees, 2008; Diana, 2010; Exare, 2017), because many words, like *ate/hate*, bear high functional load (Brown, 1988, 1991).

Conclusion

This study has attempted to explore the challenge that the consonant <h> represents for French learners of English as a foreign language. The quantitative and qualitative analyses of our recordings have revealed that L2 speech cannot be reduced to one phenomenon; despite a certain amount of intra-speaker consistency, different speech patterns emerged, ranging from the total absence of /h/ through its systematic accurate realisation to the sole presence of intrusive /h/ before initial vowels. The analysis also revealed a significant effect of gender on correct productions of /h/, with higher scores of correct realisations amongst women as well as more occurrences of intrusive /h/, where men produce glottal stops. Such an effect should be investigated further and correlated with the learners' precise language level.

While the impact of L2 misproductions on communication can be relatively strong depending on the origin of the errors (i.e., consonant, vowel, or suprasegmental) or the functional load of the target items (Brown, 1988, 1991), future studies should focus on the consequences of the /h/ phenomena under study from the point of view of native listeners. Furthermore, the correlation between learners' L2 production and L2 perception deserves to be investigated in order to determine whether–and how–English production difficulties are related to comprehension, which is as problematic for learners as target pronunciation. Finally, the parallel between such studies and L2 teaching practices should be reinforced by providing teachers with both awareness of the phenomena and tools for helping learners to overcome their difficulties.

References

ADDA-DECKER Martine, BOULA DE MAREÜIL Philippe, ADDA Gilles & LAMEL Lori, 2002, "Investigating Syllabic Structure and its Variation in Speech from French Radio Interviews", ISCA Tutorial and Research Workshop on Pronunciation Modeling and Lexicon Adaptation, Aspen Lodge, p. 89-94.

BROWN Adam, 1988, "Functional Load and the Teaching of Pronunciation", *TESOL Quarterly*, 22, p. 593-606.

BROWN Adam, 1991, *Teaching English Pronunciation: A Book of Readings*, New York, Routledge.

CAPLIEZ Marc, 2011, « Typologie des erreurs de production d'anglais des francophones : segments vs. suprasegments », *Recherches et pratiques pédagogiques en langues de spécialité*, 30, p. 44-60.

CAPLIEZ Marc & GUENDOUZ Pierre, 2021, *Apprendre et enseigner l'anglais oral*, Paris, L'Harmattan.

COLLINS Beverley & MEES Inger, 2003, Prac*tical Phonetics and Phonology: A Resource Book for Students*, London, Routledge.

DE LAUNAY Philippa, 1993, "The English 'h' Problem", *ELT Journal*, 47, p. 132-135.

DART Sarah N., 1998, "Comparing French and English Coronal Consonant Articulation", *Journal of Phonetics*, 26, p. 71-94.

DIANA Alain, 2010, « La phonétique dans l'enseignement de l'anglais aux spécialistes d'autres disciplines : enjeux et priorités », *Cahiers de l'APLIUT*, 29, p. 10-21.

ECKMAN Fred, 1977, "Markedness and the Contrastive Analysis Hypothesis", *Language Learning*, 27, p. 315-330.

ECKMAN Fred, 2008, "Typological Markedness and Second Language Phonology", *in* Jette G. Hansen Edwards & Mary L. Zampini (eds.), *Phonology and Second Language Acquisition*, Amsterdam/Philadelphia, John Benjamins, p. 95-115.

ENCREVÉ Pierre, 1988, *La liaison avec et sans enchaînement : phonologie tridimensionnelle et usages du français*, Paris, Le Seuil.

EXARE Christelle, 2017, *Les aspirations intrusives dans l'anglais des apprenants francophones*, Doctoral thesis, University of Paris 3 Sorbonne Nouvelle.

FLEGE James E., 1995, "Second Language Speech Learning: Theory, Findings, and Problems", *in* Winifred Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, Timonium, York Press, p. 233-77.

JANDA Richard D. & AUGER Julie, 1992, "Quantitative Evidence, Qualitative Hypercorrection, Sociolinguistic Variables–and French Speakers' 'eadhaches with English h/Ø", *Language & Communication*, 12, p. 195-236.

JOHN Paul & CARDOSO Walcir, 2009, "Francophone ESL learners' difficulties with English /h/", in Barbara Baptista, Andreia Rauber & Michael Watkins (eds.), *Recent Research in Second Language Phonetics/Phonology: Perception and Production*, Newcastle Upon Tyne: Cambridge Scholars, p. 118-140.

KAMIYAMA Takeki, KÜHNERT Barbara & VAISSIÈRE Jacqueline, 2011, "Do French-Speaking Learners Simply Omit the English /h/?", *Proceedings of the 17th International Congress of Phonetic Sciences*, p. 1010-1013.

MORTREUX Stéphane, 2008, *La Production des consonnes coronales en anglais par des apprenants francophones*, Doctoral thesis, University of Aix-Marseille, 303 p.

ROACH Peter, 2009, English Phonetics and Phonology–A Practical Course, Cambridge, Cambridge University Press.

VAISSIÈRE Jacqueline, 1986, "Variance and Invariance at the Word Level", *in* Joseph S. Perkell & Dennis H. Klatt (eds.), *Invariance and Variability in Speech Process*, Lawrence Erlbaum Associates, p. 534-539.

WELL John C., 1970, "Local Accents in England and Wales", *Journal of Linguistics*, 6, p. 231-252.

Annexes

Word list

hat, hand, haze, head, heat, heart, hen, hill, hollow, home, huge, hut, actor, apron, art, eat, either, ice, ill, inner, old, opera, over, umbrella.

Sentence list

Nb. The targets are underlined (initial vowels = 20; initial <h> /h/ = 16).

His name is Harry Evison.

How old is your uncle now?

Hannah asked for a hamster.

She had her hair cut in April.

I really hope they agree with us.

Emily always hates eating alone.

The author of the book is under arrest.

Hailey had a headache yesterday evening.

A horrible thing happened to Mary's husband.

We arrived at Oliver's house eight minutes after the others.

Espaces Linguistiques

Perception and production of French oral vowels by Japanese-speaking learners: does dialectal variation play a role?

Perception et production des voyelles orales du français par des apprenants japonophones : la variation dialectale a-t-elle un rôle à jouer ?

Takeki KAMIYAMA

UR TransCrit, Université Paris 8 Vincennes - Saint-Denis

takeki.kamiyama@univ-paris8.fr

Jumpei KANEDA

ATR Advanced Telecommunications Research Institute International

Ryoko HAYASHI

Kobe University

Abstract: This paper presents a series of studies on the impact of regional variation in L1 (source language) on L2 (target language) pronunciation, focusing on the high back vowel /u/ in French perceived or produced by native speakers of Japanese from the Kanto area (around Tokyo) and the Kansai area (around Osaka). 1) To serve as a base line, the 5 Japanese vowels were pronounced by 11 speakers (5 women and 6 men) from Kansai with no learning experience of French and their first 4 formants were measured to examine their acoustic properties. 2) 25 naïve listeners from Kansai took part in an AXB auditory discrimination task for French vowels including pairs /u/-/y/, /y/-/ø/ and /u/-/ø/; the results were compared with those of Kanto speakers in previous studies. 3) Some Japanese-speaking learners of French (JSL) from Kansai produced the French /u/ with a lower F2 (around 1,000 Hz for men) than typical values observed for learners from the Tokyo area. These results show limited differences for some of the speakers from Kansai compared to the tendency commonly observed with Kanto speakers.
Keywords: French vowels, Japanese speakers, regional variation, perception, production

Résumé : Cet article présente une série d'études portant sur l'impact de la variation régionale en L1 (langue source) sur la prononciation en L2 (langue cible), en se focalisant sur la voyelle fermée postérieure /u/ en français perçue ou produite par des locuteurs natifs du japonais du Kanto (région de Tokyo) et du Kansaï (région d'Osaka). 1) Pour servir de référence, les 5 voyelles du japonais ont été prononcées par 11 locuteurs (5 femmes et 6 hommes) du Kansaï (Osaka) sans expérience d'apprentissage du français et les 4 premiers formants ont été mesurés afin d'examiner leurs propriétés acoustiques. 2) 25 auditeurs originaires du Kansaï et non-apprenants du français ont effectué une tâche de discrimination auditive AXB de voyelles françaises, comprenant les paires /u/-/y/, /y/-/ø/ et /u/-/ø/ : les résultats ont été comparés avec ceux de locuteurs du Kanto dans des études antérieures. 3) Certains apprenants japonophones du français langue étrangère du Kansaï ont produit le /u/ français avec un F2 plus bas (autour de 1000 Hz pour les hommes) que les valeurs typiques observées chez des apprenants du Kanto. Ces résultats montrent des différences limitées chez certains des locuteurs du Kansaï par rapport à la tendance communément observée chez des locuteurs du Kanto.
Mots clés : voyelles françaises, japonophones, variation régionale, perception, production

Introduction

In the acquisition of second language (L2) pronunciation, the learner's source language (L1) is generally considered to have a crucial impact, as seen in influential models of L2 speech learning such as PAM-L2 (Perceptual Assimilation Model of Second Language Speech Learning: Best & Tyler, 2007) and SLM-r (revised Speech Learning Model: Flege & Bohn, 2021), as well as in more traditional accounts of Polivanov (1931), Trubetzkoy's phonological 'sieve' (1939/1969), or in contrastive analysis (Weinreich, 1953/1968; Lado, 1964). A large number of studies on L2 pronunciation or cross-language speech perception focus on a single L1 variety, often widely considered as a standard (e.g. Seoul as opposed to Gyungsang Korean or other accents).

Some, however, report cross-language studies comparing the impact of different L1 varieties on the classification of L2 vowels. Morrison (2008) ran a perceptual classification experiment using a series of synthesized stimuli covering the acoustic space acceptable to L1 English listeners as the English /ɪ/. Three groups of monolingual listeners (19 Western-Canadian English, 17 Peninsular-Spanish and 20 Mexican-Spanish speakers) were asked to classify the stimuli into phonemic categories of their L1. The stimuli for which Western-Canadian listeners' modal response was /ɪ/ were almost all identified as Spanish /e/ by Peninsular-Spanish listeners, while three-quarters of the portion of the same stimulus space was identified as Spanish /i/ and one quarter as Spanish /e/ by Mexican listeners. Chládková & Podlipský (2011) conducted a perceptual assimilation experiment of Dutch vowels by Czech listeners from Bohemia (BC: n=19) and Moravia (MC: n=22) with no knowledge of or no previous exposure to Dutch. In Bohemia, long and short high front vowels /iː/ and /i/ have spectrally different phonetic realizations ([i] and [ɪ], respectively), whereas in Moravia the spectral difference is much smaller, if not totally absent. The result of the experiment shows that BC and MC listeners perceive the Dutch vowels /i/ et /ɪ/ differently: BC labelled the tense vowel /i/ more often (61.8%) as the Czech long vowel /iː/ than MC did (34.4%), and the difference turns out to be significant (p < .001).

These results, obtained in cross-language studies with non-learners, suggest that the variety of L1 could also have a considerable impact on the acquisition of phonemic contrasts amongst L2 learners.

Based on the above-mentioned literature, the goal of the present paper is to present a series of studies enabling a comparison of the influence of two L1 regional varieties of Japanese on the perception and production of the high back vowel /u/ in French by native speakers of Japanese,

contrasted by neighbouring vowels. The L1 varieties under scrutiny are 1) from the greater Tokyo area (Kanto region), often considered to be a "national standard" (Shibatani, 1990), and 2) from the Kansai region, including Osaka (Figure 1).

Figure 1: Kansai (including cities of Osaka, Kyoto, Kobe and Nara) and Kanto (including Tokyo) regions.
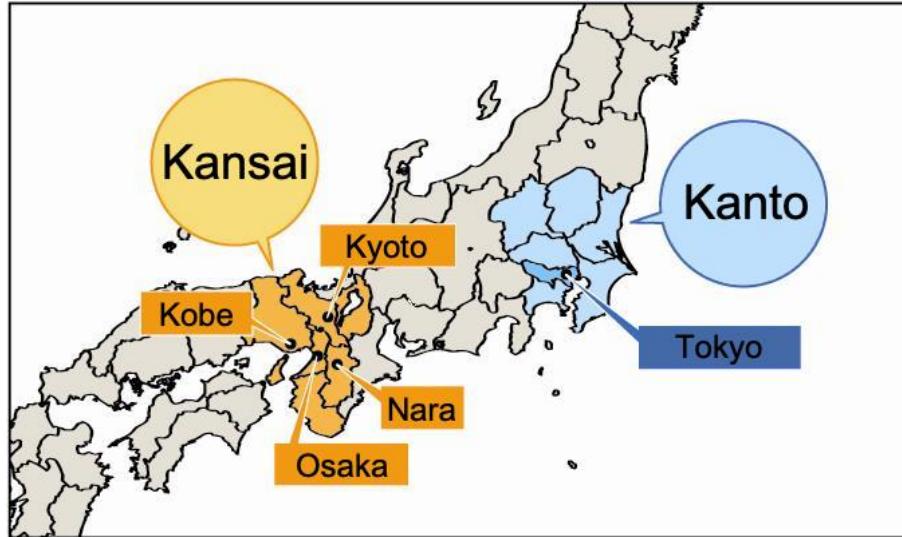
Figure adapted from Tanaka, Tetsuya "*Nihongo kyouiku tsûshin: jugyou no hinto: hougen (kansaiben) ni fureru* (Japanese Language Education Newsletter: Lesson tips: Exposure to dialects - Kansai dialect)", Japan Foundation[38].

Figure 2: Vowel inventories of Reference French (left: based on Vaissière, 2006, *inter alia*) and of Tokyo (Kanto) Japanese (right: based on Shibatani, 1990, Sugitô, 1995, Vance, 2008, *inter alia*).

---

[38] https://www.jpf.go.jp/j/project/japanese/teach/tsushin/hint/201107.html

The target variety, Reference French (RF), as described by Detey, Lyche *et al.* (2016)[39], was chosen for two reasons. First, the learners who participated in the experiment resided in Japan and had very little experience abroad in a French-speaking area. Second, the RF variety is widely adopted in teaching material used in foreign language settings. RF has 3 nasal vowels, and 10 oral vowels, which include a series of front rounded vowels /y/, /ø/ and /œ/. By contrast, the 5-vowel system of Tokyo and Osaka Japanese does not include such front rounded vowels (Figure 2). These vowels, together with /u/, can be a challenge for Japanese L1 learners of French (JSL) (Kamiyama, 2011). While both languages have a phoneme commonly transcribed as /u/, the high back vowel in RF is realized as a focal vowel (Schwartz et al., 1997) with F1 and F2 being close together and under 1,000 Hz, thereby forming a frequency zone of high concentration of energy (Liénard, 1977; Vaissière, 2007, *inter alia*). It has been shown that native speakers of Tokyo (Kanto) Japanese who learn French as a foreign language (FFL) tend to produce this target vowel /u/ with a higher F2 without a high concentration of energy for F1 and F2 (Figure 3). This corresponds to a tongue position typically more fronted than that of native speakers (Kocjančič Antolík, Pillot-Loiseau & Kamiyama, 2019). The vowel tokens thus produced by JSL may be perceived by native listeners of French as the mid-high front rounded /ø/, characterized by evenly distributed formants without a zone of high concentration of energy (Kamiyama & Vaissière, 2009). In perception, the vowel contrast /u/-/ø/ is among the most difficult for JSL to distinguish (Kamiyama & Vaissière, 2009), as expected from French loanwords in Japanese, where both vowels are usually adapted to /u/ in Japanese ("Strasbourg" /stʀasbuʀ/ -> ストラスブール /sutorasubuRru/; "Périgueux" /peʀigø/ -> ペリグー /periguR/[40]), suggesting the case of Single-Category (SC) assimilation (or Category-Goodness assimilation) in PAM (Perceptual Assimilation Model: Best, 1995). A perceptual categorization experiment using stimuli made by articulatory synthesis with Maeda's articulatory parameters (Maeda, 1982) corroborates the fact that the acoustic-articulatory space of French native listeners' /u/ and /ø/ corresponds roughly to the acoustic-articulatory space of /u/ for Japanese native listeners (Kamiyama, 2011).

---

[39] The authors prefer this term to that of *Standard French*, "seen by many as too prescriptive", but at the same time, they acknowledge that "unfortunately, all too often, Reference French represents exactly the same reality: an abstraction, a set of features attributed to a certain social class (educated people) and geographical area (Paris and surroundings)" (Detey, Lyche *et al.*, 2016, p. 58-59).

[40] /R/ is a phoneme that is realized as the lengthening of the preceding vowel in Japanese (cf. Labrune, 2012; some authors use /H/ instead of /R/: Vance, 2008).

Figure 3: Wide-band spectrogram (window length: 5 milliseconds), obtained using Praat (Boersma & Weenink, 2007), of French /u/ in isolation pronounced by a male native speaker of French (left), and by a male JSL from Tokyo (right).
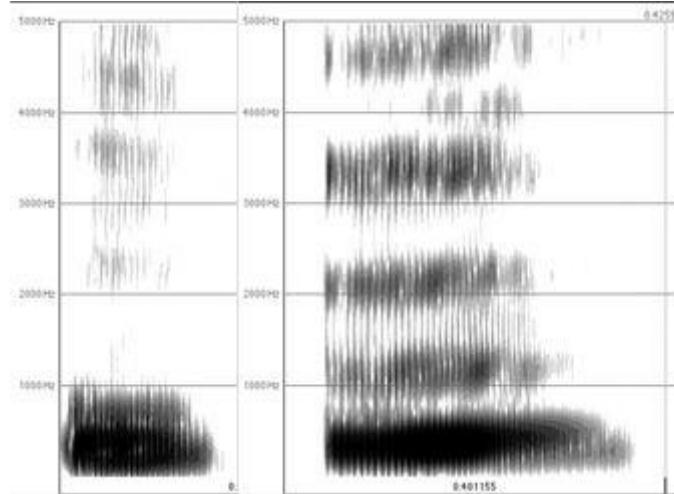
The trend described above stems from the fact that the high non-front vowel in Kanto (Tokyo) Japanese /u/ shows a higher F2 (> 1,000 Hz) than that of the French /u/. By contrast, the vowel /u/ in Kansai Japanese is usually described as more rounded (Shibatani, 1990) and its F2, as well as F1, is lower (Sugitô, 1995) than that of Kanto (Figure 4).

Figure 4: F1 (vertical axis) and F2 (horizontal axis) of the 5 vowels in Osaka (Kansai: left) and Tokyo (Kanto: right) Japanese pronounced by male speakers.
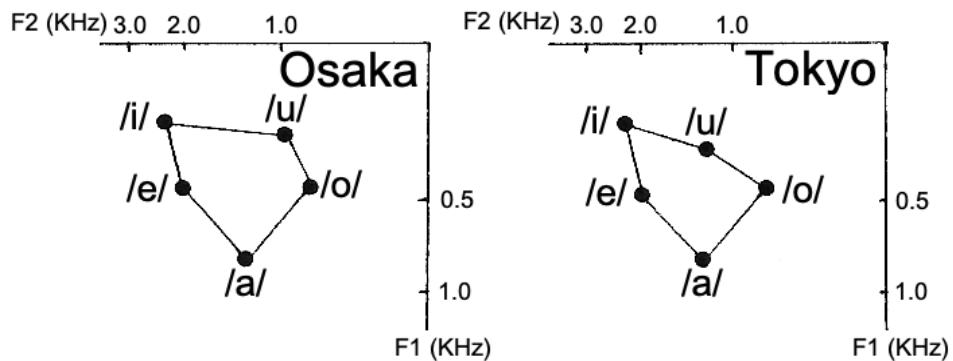
Figures adapted from Sugito (1995)

The articulatory and acoustic characteristics of the vowel /u/ in Tokyo (Kanto) Japanese are described as follows: /u/ in isolation shows a more front, and especially lower tongue position, with less protruded lips (Kokuritsu Kokugo Kenkyûjo, 1990; Takebayashi, 1996) compared to the French /u/ (Bothorel *et al.*, 1986); acoustically, F2 of /u/ varies mostly between 1,000 and 1,500 Hz

for male speakers in various consonantal contexts, which is comparable to F2 of /a/ among other vowels (Figure 5, adapted from Mokhtari & Tanaka, 2000). Little is known, however, about articulatory and acoustic properties of Kansai Japanese vowels, except for the illustrative descriptions mentioned above. It is therefore unclear whether Japanese listeners from Kansai may be influenced by possibly low values of F2 in /u/ and categorize non-native vowels differently from Kanto listeners, or whether they also benefit from these articulatory and acoustic properties when they acquire second or foreign language vowels (PAM and PAM-L2: Best, 1995; Best & Tyler, 2007).

Figure 5 : F1 (horizontal axis) and F2 (vertical axis) of the 5 vowels in Tokyo (Kanto) Japanese pronounced by 5 male speakers. 22 words containing either a long vowel /VR/ or a hiatus /VV/
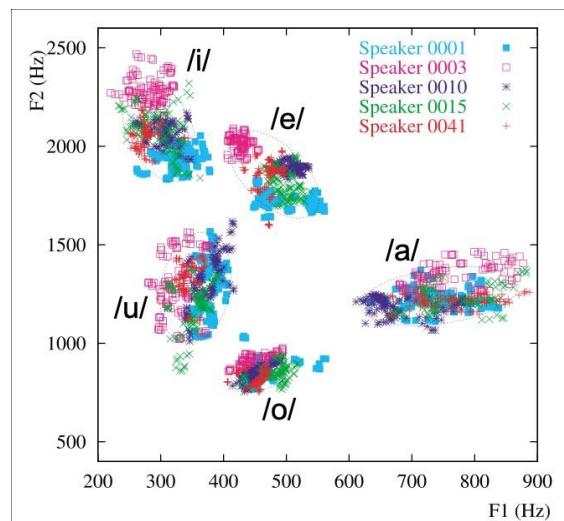
Figure adapted from Mokhtari & Tanaka (2000)

The findings presented above lead us to the following hypotheses:

1) Kansai-Japanese speakers produce /u/ with lower F2 than Kanto speakers in their L1 Japanese.

2) Kansai-Japanese listeners better distinguish the French /u/ from other French vowels which may sound similar for Japanese listeners in general, than Kanto listeners.

3) Kansai-Japanese speakers learning French as a foreign language learn to produce more native-like tokens of the French /u/ than learners from Kanto.

These hypotheses motivated three experiments, presented in the following sections:

1) Production of the 5 vowels in Japanese by native speakers from Kansai

2) AXB discrimination task for French vowel pairs by naive listeners (non-learners of French) from Kansai

3) Production of French vowels by Japanese-speaking learners of French from Kansai.

1. Experiment 1: acoustic analysis of Kansai Japanese vowels

In the first experiment, the acoustic properties of the 5 Japanese vowels produced by Kansai speakers were examined to observe the actual tendency amongst speakers and to serve as a baseline.

1.1. Method

11 speakers (5 women and 6 men) from Kansai (Hyogo, Osaka, Nara, Shiga, Wakayama prefectures), aged 18 to 23, all students at Kobe University, participated in this experiment. One of them spent 3 years from age 6 to 8 in Ibaraki Prefecture in Kanto, and another lived in Germany for a year at the age of 21, but none of the others had lived outside Kansai. Their self-assessed degree of use of *kyoutsûgo* (literally, "common language", a term referring to non-dialectal variety, often considered close to Tokyo Japanese) ranged from 1 (not frequent) to 5 (frequent). The 5 Japanese vowels in isolation /i/ /e/ /a/ /o/ /u/, as well as the sequences /ja/ /ju/ /jo/ were embedded in the carrier sentence /sore o _____ to iu/ ("one calls that …") presented one by one on a screen in a semi-random order. The test items were preceded by 3 training sentences in colloquial Kansai (Osaka) Japanese taken from Sugito (1995). The list of carrier sentences was read 5 times. The participants wore a head-set microphone and the audio data were recorded at 44.1 kHz, 16 bits, using ROCme! (Ferragne *et al.*, 2012). The procedures were conducted by a male researcher (from Tokyo) the participants met for the first time. The first 4 formants were measured in three zones, as mean values during the first, second, and last third of the vowel portion. It should be noted that no clear diphthongization tendency has been reported about Japanese vowels in isolation. This was also the case in the present data set.

1.2. Results

The mean F2 of /u/, shown in Figure 6, mostly ranges between 1,000 Hz and 1,500 Hz for women and 1,000 Hz and 1,300 Hz for men, which is comparable to that of /a/.

Figure 6: The mean values and Standard Deviation (SD) of the first 4 formants of the 5 Japanese

vowels /i/ /e/ /a/ /o/ /u/ pronounced by 5 women (left) and 6 men (right) from Kansai

(3 measures per token, 5 repetitions). The error bars represent ±1SD
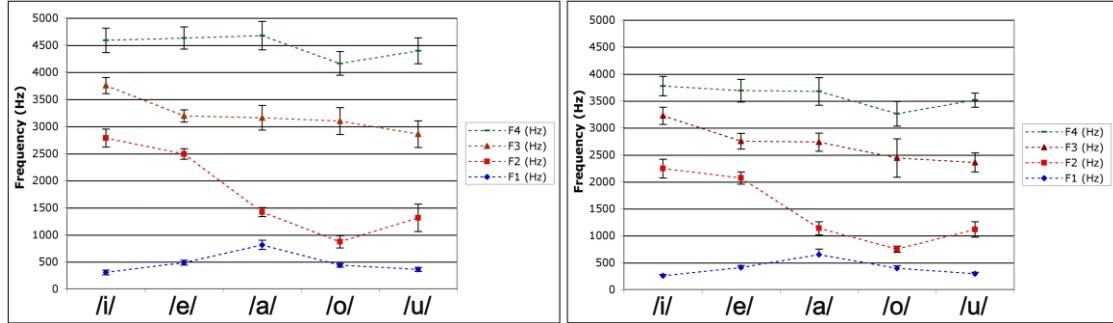
Taking a closer look at the distance between F1 and F2, which is known to be larger in the Kanto

(Tokyo) Japanese /u/ than the French /u/, reveals a higher inter-speaker variability than for the

other non-front vowels /o/ and /a/, as shown in Figure 7. The relatively low F1-F2 distance for

the female speaker sp08 and the male speakers sp02, sp04, sp05 and sp06 is caused by F2 ranging

between 1,000 and 1,100 Hz. Some speakers showed a higher intra-speaker variability than others:

for sp04, for example, mean F2 ranged between 1,150 and 1,200 Hz in the first two repetitions,

but then dropped below 1,000 Hz in the other three repetitions.

F1-F2 distance tends to be higher for speakers who gave a higher rating to the self-assessed degree

of use of the non-local form, "common language" (*kyoutsûgo*), even though this tendency is not

systematic.

Figure 7: The mean values and Standard Deviation (SD) of the distance between F1 and F2 of

the Japanese vowel /u/ pronounced by 5 women (left) and 6 men (right) from Kansai

(3 measures per token, 5 repetitions). The number after Speaker ID represents the self-assessed

degree of use of the "common language" (k*youtsûgo).* The error bars represent ±1SD
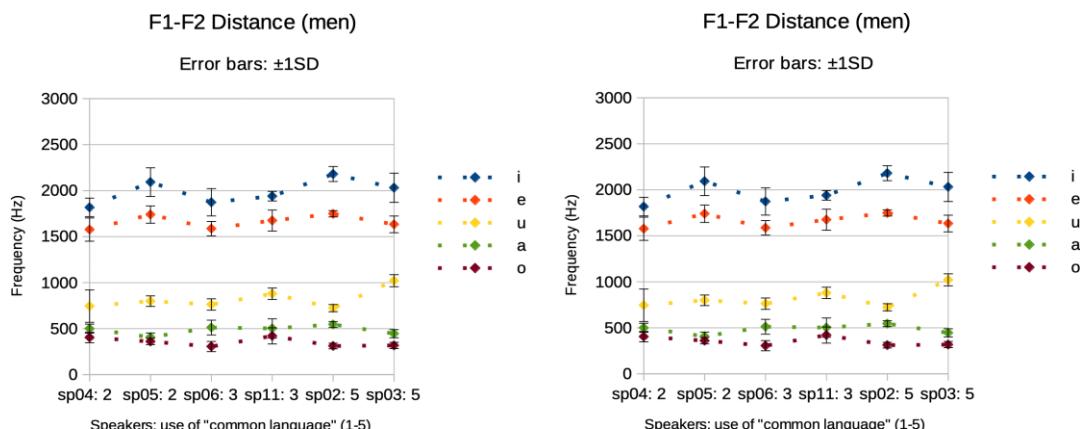
The findings suggest that Kansai listeners are indeed exposed to /u/ tokens with relatively low F2 to some extent at least (compared to Kanto listeners), in spite of variability, which might impact the perception of the French /u/, produced typically with even lower F2 in isolation.

2. Experiment 2: Auditory discrimination of French vowel pairs by naïve listeners from Kansai

Following the results of Experiment 1, an AXB auditory discrimination task was set up to examine the perceptual distinction of French vowel pairs including /u/ by naïve listeners of Kansai Japanese.

2.1. Method

The participants were 25 students enrolled in either of two different universities located in the cities of Osaka and Kobe. None of them had learnt French or any other language characterized by having a series of front rounded vowels in its vowel inventory.

Six vowel pairs /u/-/y/, /y/-/ø/, /u/-/ø/, as well as /i/-/e/, /u/-/o/ and /ɛ/-/a/, were used to compose 72 triplets of vowels in isolation: 6 pairs x 4 combinations and orders of tokens (/u u y/, /u y y/, /y u u/ and /y y u/ for the vowel pair /u/-/y/) x 3 speaker conditions: 1) all 3 stimuli in the AXB triplet were pronounced by the same female native speaker of Reference French; 2) "A" (the first stimulus in the triplet) and "B" (the last) were pronounced by the same female native speaker as in the previous condition, but "X" (the second one) was pronounced by another RF female native speaker; 3) "X" was pronounced by a male RF native speaker, while "A" and "B" were produced by the same female speaker as in the first speaker condition. The vowel contrasts /i/-/e/ and /o/-/u/ were included since some cases of incorrect identification were observed in a perceptual identification test administered for FFL in Kamiyama (2011). By contrast, it is predicted from the results of the same study that /ɛ/-/a/ will be discriminated almost perfectly. The mean duration of the vowel tokens was 180 milliseconds (ms). The intra-stimulus interval (between the stimuli in each triplet) was set to 1 second, so that the stimuli would be processed as linguistic (phonemic) units rather than physical (acoustic) ones. The list of 72 triplets were played 2 times in different orders. The experiment was self-paced and conducted using Praat (Boersma & Weenink, 2007).

2.2. Results

The percentage correct is represented in Figure 8. Among the pairs including the vowels /u/, /y/ and /ø/, the contrast /u/-/ø/ was the most difficult to distinguish perceptually (78.3 % correct), followed by /y/-/ø/ (83,4% correct) and /u/-/y/ (93.4% correct). This order is the same as for listeners from Kantô: for 7 non-learners of French, 76.2% correct for /u/-/ø/, 79.8% correct for /y/-/ø/, 85,7% correct for /u/-/y/; for 14 learners of French as a foreign language, 84.8% correct for /u/-/ø/, 91.1% correct for /y/-/ø/, 94.9% correct for /u/-/y/ (Kamiyama & Vaissière, 2009).

Figure 8: Mean percentage correct of discrimination for French vowel pairs perceived by 25 non-learners of French from Kansai. 2 repetitions x 12 triplets x 6 vowel pairs. Chance level: 50%.
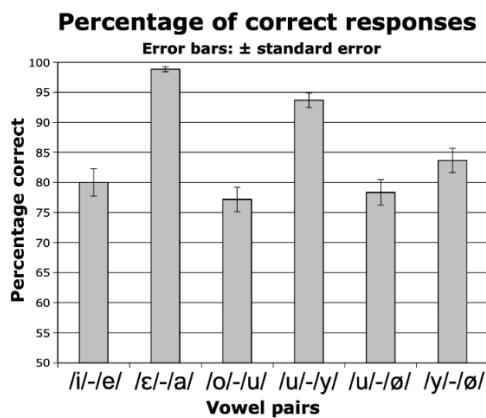
3. Experiment 3: Production of French vowels by JSL from Kansai

The previous experiment shows that the French phonemic contrast /u/-/ø/ is as difficult for Kansai listeners to distinguish perceptually as for Kanto listeners. The present study sheds light on the production of these two vowels among other French vowels by learners of French as a foreign language from Kansai in comparison with those from Kanto.

3.1. Method

The speech material (part of the corpus reported in Landron *et al.*, 2010) consists of the 13 RF vowels (10 oral /i e ɛ a ɔ o u y ø œ/ and 3 nasal (/ɛ̃ ɑ̃ ɔ̃/), embedded in carrier sentences such as "*Bébé, je dis « é » comme dans bébé*" (Baby, I say "é" as in baby). In the present study only the 10 oral vowels are analyzed. The target vowel was elicited with the help of a keyword (e.g. "*bébé*", "*pouce*", "*deux*" containing /e/, /u/, /ø/, respectively) and a typical spelling pattern (e.g. « é » /e/, « ou » /u/, « eu » /ø/).

Two subgroups of JSL of FFL participated in the experiment. The first consisted of 4 students (2 women and 2 men) enrolled at Tokyo University of Foreign Studies. They had all lived in the Greater Tokyo area (Kanto) for at least 3 or 4 years, but originally came from various regions, including one participant from Tokyo, one from Kagawa, located across the Seto Inland Sea from Kansai. Two of the participants had started learning French at 17-18 years old and the others earlier (13 and 15 years old), which makes 3 to 9 years of learning experience. The second subgroup was composed of 4 second- and third-year students (2 women and 2 men) at Kobe University (Hyogo Prefecture), approximately 30 km West of Osaka. They all grew up mainly in Kansai. They had all started learning French at university (approximately at 18 years old), which makes 1.5 to 2.5 years of learning experience. All 8 participants were in the age range of 18-24 years.

The carrier sentences were presented one by one in a semi-random order on a computer screen. The sentence list was repeated 4 times without a break. The participants were invited to separate the target vowel from the rest of the carrier sentence so as to avoid formant transitions due to the adjacent consonants as far as possible. A training session preceded the recording session. The recording took place in recording studios of Tokyo University of Foreign Studies and Kobe University (Faculty of Global Human Sciences: named "Faculty of Intercultural Studies" at the time of the recording) using a headset microphone. The learners' production was recorded at 44.1 kHz, 16 bits. The target vowels were annotated manually using Praat (Boersma & Weenink, 2007). Portions where F2 and higher formants are not clearly observable, or where irregular periods are found, were excluded. The formant values were extracted every 6.25 milliseconds and the mean value was taken for the first, second, and last thirds of the vowel portion. The automatic formant detection was checked and the parameters were modified when necessary.

3.2. Results

The formant values of the vowels produced by the 4 learners from Kanto are presented in Figure 9.

Figure 9: Learners from Kanto: mean F1, F2, F3 and F4 of the French oral vowels
(3 measures x 4 repetitions), by 4 learners (2 women above; 2 men below). /e/ produced 3 times
by the learner 2ME. Error bars: ± 1SD

Learner 1FE

	i	e	ε	a	ɔ	o	u	y	ø	œ
F4	4356	4300	4351	4090	4239	4145	4410	4434	4385	4408
F3	3193	2866	2786	2603	2726	2757	2443	2545	2489	2463
F2	2731	2296	2224	1292	839	845	1447	1524	1548	1572
F1	367	564	551	774	510	524	362	375	383	372

Learner 3FE

	i	e	ε	a	ɔ	o	u	y	ø	œ
F4	5351	5518	5529	4567	3971	3944	4131	4048	4070	4186
F3	3614	3368	3353	2884	2696	2669	2708	2653	2696	2776
F2	2839	2549	2514	1335	837	859	1829	1844	1995	2109
F1	353	537	576	864	525	531	352	360	360	370

Learner 2ME

	i	e	ε	a	ɔ	o	u	y	ø	œ
F4	4195	4299	4216	3857	3951	4107	3647	3346	3329	3560
F3	3356	3120	3081	2694	2604	2678	2482	2261	2348	2388
F2	2338	2426	2423	1456	675	601	829	2048	1672	1652
F1	323	421	435	895	417	404	396	258	378	419

Learner 4ME

	i	e	ε	a	ɔ	o	u	y	ø	œ
F4	3643	3697	3678	3762	3345	3313	3375	3390	3387	3371
F3	3311	2952	2762	2554	2222	2245	2029	2067	2167	2139
F2	2108	2085	1961	1284	813	745	1329	1496	1598	1611
F1	383	458	510	820	482	437	310	348	401	418

It was decided not to limit the graphic representation to rounded vowels but to include all oral vowels, for it is important to observe vowels as a whole system of contrasts. Three of the learners (1FE, 3FE and 4ME) show F2 values of /u/ higher than 1,400 Hz and far from F1. This tendency corroborates the observation made in previous studies on Kanto learners (Kamiyama & Vaissière, 2009). The only participant who produced the first two formants close together under 1,000 Hz (2ME) had started learning French earlier than the other participants, at the age of 13. His /y/ is also realized with close F2/F3 (Liénard, 1977, among others) with a small Standard Deviation (SD), which indicates that this vowel is not diphthongized either, unlike cases observed for some JSL (falling F2 due to an onglide [j], like in the Japanese sequence /ju/: Kamiyama & Vaissière, 2009).

Let us note that formant values are generally higher for women than men, but to a lesser extent for the formants essentially due to the Helmholtz resonance (F1 of /i/ /y/ /u/ and F2 of /u/ in French: Tubach, 1989).

Figure 10: Learners from Kansai: mean F1, F2, F3 and F4 of the French oral vowels

(3 measures x 4 repetitions), by 4 learners (2 women above; 2 men below). Error bars: ± 1SD

Learner 5FW

	i	e	ɛ	a	ɔ	o	u	y	ø	œ
F4	4536	4638	4474	4355	4345	4197	4422	4383	4390	4383
F3	3518	3314	3210	3108	3081	3016	2918	2893	2879	2922
F2	2875	2383	2440	1321	988	945	1450	1526	1559	1459
F1	352	564	551	779	515	507	400	374	387	399

Learner 6FW

	i	e	ɛ	a	ɔ	o	u	y	ø	œ
F4	4557	4444	4518	4124	4217	4017	4378	4249	4310	4350
F3	3827	3424	3363	3409	3352	3313	3208	3184	3184	3170
F2	2937	2528	2467	1443	943	878	1521	1653	1644	1677
F1	314	562	558	842	443	447	459	445	443	439

Learner 7MW

	i	e	ɛ	a	ɔ	o	u	y	ø	œ
F4	3638	3532	3517	3283	3135	3195	3304	3265	3347	3301
F3	3221	2708	2710	2646	2444	2458	2258	2256	2264	2251
F2	2290	2117	2043	1153	691	659	920	1876	841	887
F1	231	405	423	606	394	376	260	233	248	255

Learner 8MW

	i	e	ɛ	a	ɔ	o	u	y	ø	œ
F4	4031	3549	3510	3648	3621	3611	3484	3459	3636	3633
F3	3218	2814	2728	2728	2585	2554	2377	2171	2448	2529
F2	2308	2182	2094	1254	750	687	1035	1881	1510	1442
F1	228	330	360	741	317	324	230	225	344	411

The results of the Kansai learners are presented in Figure 10. Concerning /u/, the two women (5FW and 6FW) show F2 around 1,500 Hz like the learners from Kanto. By contrast, the two men (7MW and 8MW) present F2 around 1,000 Hz (920 Hz for 7MW and 1,035 Hz for 8MW; still significantly higher than the native speakers of French reported in Kamiyama & Vaissière, 2009: Wilcoxon rank sum test, $p < .0001$), even if the F1-F2 distance is relatively large compared to native speakers: 660 Hz for 7MW, 805 Hz for 8MW, against 342 Hz and 333 Hz for two male native speakers in Kamiyama & Vaissière (2009), the difference being significant (Wilcoxon rank sum test, $p < .0001$).

In addition to the phonetic realization of /u/, the distinction of this vowel from other target vowels was also examined by means of Euclidean distance. This measure was already applied to the data of French rounded vowels pronounced by JSL in order to estimate the degree of dispersion depending on the speaking tasks adopted (Marushima *et al.*, 2010: calculation based on the first two formants). In the present study, the Euclidean distance was obtained from the first four formants in Bark scale ([26.81/(1+1960/f)]−0.53: Traunmüller, 1990) so that the perceptual aspect would be taken into account [1]:

[1] Euclidean distance between vowels A and B

$$= \sqrt{(F1_{(A)} - F1_{(B)})^2 + (F2_{(A)} - F2_{(B)})^2 + (F3_{(A)} - F3_{(B)})^2 + (F4_{(A)} - F4_{(B)})^2}$$

Figure 11 represents the Euclidean distances between /u/-/o/, /u/-/y/, /y/-/ø/ et /u/-/ø/, in comparison with the values observed for native speakers of French who produced vowels in isolation in a similar carrier sentence (Kamiyama & Vaissière, 2009). The learners' (from Kanto and Kansai) formant values distinguish /u/ and /o/ as clearly as native speakers, which is not always the case with the other contrasts. It was noted earlier that the learner 7MW from Kansai pronounced /u/ with F2 lower than 1,000 Hz, but the small Euclidean distance between /u/-/ø/ seen in this figure shows that these two target vowels were produced in a similar manner.

Figure 11: Euclidean distance (based on mean F1, F2, F3 and F4 in Bark scale) between /o/-/u/, /u/-/y/, /y/-/ø/ and /u/-/ø/: women (left) and men (right): learners from Kanto (FE, ME) and Kansai (FW, MW); native speakers of French (fr) in Kamiyama & Vaissière (2009). 4 repetitions for learners, 3 for native speakers.
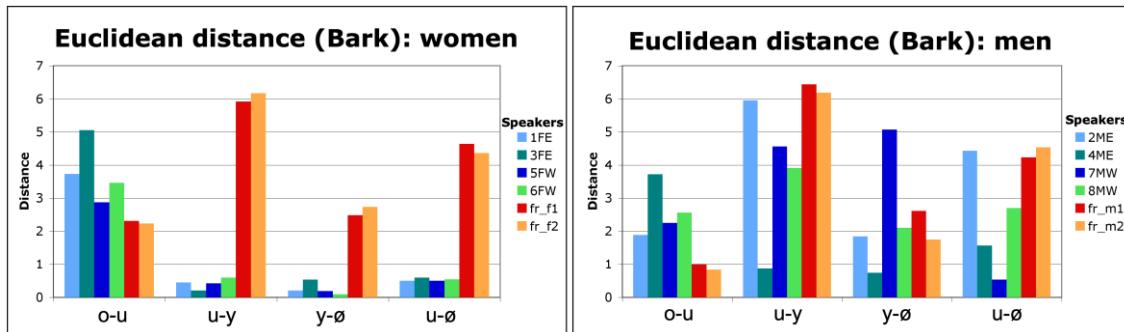
3.3. Discussion

The results of the series of experiments presented in the previous sections lead us to reconsider the hypotheses suggested earlier.

For Experiment 1, it was hypothesized that Kansai-Japanese speakers do indeed produce /u/ with lower F2 than Kanto speakers in their L1 Japanese. This hypothesis was partly adopted: for 5 of the 11 speakers, mean F2 ranged between 1,000 and 1,100 Hz, and an another showed a high intra-speaker variability: between 1,150 and 1,200 Hz in the first two repetitions, and then below 1,000 Hz in the three other repetitions. In examining the inter- and intra-speaker variability observed in this local accent not considered as a national standard, sociolinguistic factors need to be taken into account. The participants were university students and the recording took place in

their university, with an investigator coming from Tokyo, whom they met for the first time. All of these factors might have favoured a more formal speaking style than, for example, that used in an informal and casual speech with family members or childhood friends. Different tasks and settings could elicit context-dependent variability, as designed for the PFC (*Phonologie du français contemporain*: Phonology of Contemporary French) protocol, with word lists, text reading, formal and informal conversations (Detey, Durand *et al.*, 2016). The self-assessed degree of the use of *kyoutsûgo* (common language) showed a slight tendency to favour higher F2 for those who use it more frequently: other sociolinguistic factors such as the age of the speakers, the gender, the socio-economic status, the interlocutor (e.g. peer speaking in the same accent variety or external investigator using a stardardized variety) and the setting of the data elicitation (speaking task) might help clarify further the tendency observed in the present dataset.

The current study and the previous ones cited refer to articulation but only indirectly or with simple observations: to the best of our knowledge, there is no articulatory data available on the production of Japanese vowels in other varieties than Tokyo (Kanto) Japanese (e.g. Kokuristu Kokugo Kenkyûjo, 1990), except for the data from 5 Kansai speakers in the recently published Real-time MRI Articulatory Database (Version 1)[41] presented in Maekawa (2023). Amongst the 5 Kansai speakers, born between 1952 and 1970, who produced the sequence /uR/ (long /u/) in this database, 4 of them present relatively low F2 (< 1,100 Hz), corroborating previous descriptions and our findings. Furthermore, the female speaker born in 1952 produced F2 values as low as those of French native speakers' /u/ (c. 600 Hz). A qualitative observation of their tongue position and the lip configuration seems to reveal the following: 3 speakers (5 tokens) out of 5 (7 tokens) from Kansai are characterized by a considerable degree of lip protrusion, whereas only 3 "standard" speakers (5 tokens) out of 15 (21 tokens) showed the same articulatory configuration; the tongue position is considerably back for 3 speakers (4 tokens) out of 5 (7 tokens) from Kansai and is approximately central for the other 2 speakers (3 tokens), while it is considerably back for 1 "standard" speaker (1 token) only out of 15 (21 tokens), relatively back for 7 speakers (9 tokens), and approximately central for the other 7 speakers (10 tokens); all 5 Kansai speakers tend to share narrower constriction areas both for the lips and the tongue compared to Kanto speakers, contributing to lower both F1 and F2. Quantitative analyses of these articulatory data would provide further insights into the production of the Kansai /u/, among other vowels, in comparison with Kanto speakers.

---

[41] https://rtmridb.ninjal.ac.jp

The AXB auditory discrimination test in Experiment 2 revealed that the order of difficulty for the contrasts /u/-/ø/ (the most difficult), /y/-/ø/ and /u/-/y/ (least difficult) was the same for non-learners of French from Kansai as for those from Kanto and learners from Kanto. This finding does not support Hypothesis 2: Kansai-Japanese listeners distinguish the French /u/ perceptually from other vowels sounding similar to Japanese listeners more easily than Kanto listeners, thanks to their exposure to the Kansai /u/, which is acoustically closer to the French /u/ than the Kanto (Tokyo) Japanese /u/ is. Even though the Kansai listeners showed a slightly higher percentage correct for the contrasts /u/-/y/ and /y/-/ø/ than the Kanto listeners, a more comparable dataset will be needed to estimate the impact of the differences observed.

Compared with the previous cross-language studies on classification to L1 categories (Morrison, 2008; Chládková & Podlipský, 2011), it turns out that the vowel contrast /u/-/ø/ is not facilitated by a different dialectal background in the auditory discrimination task. As mentioned earlier, this contrast could be considered as a case of Single-Category (CS) in PAM (Best, 1995). In this model, Category-Goodness (CG) pattern, in which one of the L2 phonemes corresponds to a better token of an L1 category than the other L2 phoneme, perceived as a less good token of the same L1 category, predicts a better discrimination than in CS. If the French /ø/ were perceived as a less good token of the Japanese /u/ than the French /u/ were by Kansai listeners, unlike Kanto listeners, then the discrimination of the contrast /u/-/ø/ would be better for the former group than the latter. The fact that it is not the case suggests not only that there is no phonetic category (allophone) for the Kansai-type [u] with lower F2, apart from the Kanto-type [u] with higher F2, but also that these two types form a continuum with neither of them as a prototype, probably due to the intra- and inter-speaker variability, which, together with exposure to the Kanto variety through media and communication with speakers from Kanto and other regions, makes the Kanto-type [u] a frequently observed type of token.

It was shown in Experiment 3 that the two male learners from Kansai produced the French /u/ with F2 around 1,000 Hz (one of them lower than 1,000 Hz), but with a larger F1-F2 distance than the native speakers, and that the participant who produced F2 lower than 1,000 Hz also produced /ø/ in a similar manner. Let us note that the small distance between F1 and F2 characterizes the French /u/ (Gendrot *et al.*, 2008) and that the higher formants, including F3, of the French back vowels /u/ and /o/ are not perceptible (F'2 close to F2: Vaissière, 2011). This

result suggests that the phonetic realization of /u/ in Kansai Japanese facilitates the phonetic acquisition of the French /u/ to some extent, even if the first two formants are not as close as in native speakers' tokens. An auditory assessment test with native-speaking listeners from two dialectal backgrounds (RF speakers from Paris and speakers from Québec) shows that the RF listeners examined perceived those stimuli of /u/ (produced by JSL from Kanto) with F2 between 1,000 and 1,100 Hz as /ø/ and /u/ almost equally often, but considered as very poor exemplars of either of them, while the Québec French listeners tested identified the same stimuli as /u/ in the majority of the cases, with a better goodness rating than RF listeners (Tremblay & Kamiyama, 2009). This result suggests that the tokens produced by the two male learners from Kansai are also accepted quite well as /u/ by Québec listeners, since their F2 is located between 1,000 and 1,100 Hz. A more systematic study including Kansai learners would clarify the perceptual impact of this type of phonetic realization on native-speaking listeners' assessment. By contrast, the phonemic acquisition of the production of /u/, namely learning to pronounce it distinctly from neighbouring vowels, especially /ø/, is not necessarily facilitated, as shown by the Euclidean distance observed for the vowel contrast /u/-/ø/.

Considering these findings, Hypothesis 3, stating that Kansai-Japanese speakers learning French as a foreign language learn to produce more native-like tokens of the French /u/ than learners from Kanto, is accepted literally, but cannot be extended to phonemic acquisition of the vowel phoneme: the phonemic contrast between /u/ and other vowels, especially /ø/, is still part of the difficulties. A further study with a larger number of participants would enable an observation of more general tendencies.

The present series of experiments examined the vowels in isolation, but it will be also necessary to study different consonantal and prosodic contexts in further studies. Indeed, Gendrot and Adda-Decker's (2005) data present higher F2 for the French /u/ (1,153 Hz for women), which is probably due to the effect of coarticulation in continuous speech, where vowels are located in various consonantal and prosodic contexts. Likewise, F2 of /u/ in (Kanto - Tokyo) Japanese is also variable in various consonantal contexts, as shown in the data presented by Mokhtari and Tanaka (2000), cited earlier in Figure 5.

Conclusion

This paper reported a series of studies on the possible impact of the L1 varieties, namely, the Japanese varieties of Kansai and Kanto, on the acquisition of the French high back vowel /u/ in contrast with some neighbouring vowels. In spite of intra- and inter-speaker variability, some speakers from Kansai produce the Japanese /u/ with lower F2 than typical Kanto Japanese values, but this tendency does not seem to help them to better distinguish the French /u/-/ø/ perceptually. When learners produce the French vowel /u/, the Kansai-type [u] with lower F2 may facilitate the phonetic realization of the target vowel, but it does not necessarily mean that the phonemic contrast /u/-/ø/ is also acquired. These findings suggest that being aware of the regional or individual differences of learners may be useful in teaching foreign or second language pronunciation.

References

BEST Catherine T., 1995, "A direct realist view of cross-language speech perception", *in* Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, Baltimore, York Press, p. 171-204.

BEST Catherine T. & TYLER Michael D., 2007, "Nonnative and second-language speechperception: Commonalities and complementarities", *in* Murray J. Munro & Ocke-Schwen Bohn (eds.), *Language experience in second language speech learning: In honor of James Emil Flege*, Amsterdam, John Benjamins, p. 13-34.

BOERSMA Paul & WEENINK David, 2007, *Praat: doing phonetics by computer* (Version 4.6.13) [Computer program]. Retrieved in August 2007, from http://www.praat.org/.

BOTHOREL André, SIMON Péla, WIOLAND François & ZERLING Jean-Pierre, 1986, *Cinéradiographie des voyelles et consonnes du français*, Strasbourg, Institut de Phonétique de Strasbourg.

CHLÁDKOVÁ Kateřina & PODLIPSKÝ Václav Jonáš, 2011, "Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners", *The Journal of the Acoustical Society of America*, vol. 130(4), p. EL186–EL192.

DETEY Sylvain, DURAND Jacques, LAKS Bernard & CHANTAL Chantal, 2016, "The PFC programme and its methodological framework", *in* Sylvain Detey, Jacques Durand, Bernard Laks & Chantal Lyche (eds.), *Varieties of Spoken French*, Oxford, Oxford University Press, p. 13-23.

DETEY Sylvain, LYCHE Chantal, RACINE Isabelle, SCHWAB Sandra & LE GAC David, 2016, "The notion of norm in spoken French: Production and perception", *in* Sylvain Detey, Jacques Durand, Bernard Laks & Chantal Lyche (eds.), *Varieties of Spoken French*, Oxford, Oxford University Press, p. 55-67.

FERRAGNE Emmanuel, FLAVIER Sébastien & FRESSARD Christian, 2012, *ROCme!* (Version 2.0) [Computer software], retrieved June 4, 2012, Available from www.ddl.cnrs.fr/rocme.

FLEGE James Emil & BOHN Ocke-Schwen, 2021, "The Revised Speech Learning Model (SLM-r)", *in* Ratree Wayland (ed.), *Second Language Speech Learning: Theoretical and Empirical Progress*, Cambridge, Cambridge University Press, p. 3-83. doi.org/10.1017/9781108886901.002.

GENDROT Cédric & ADDA-DECKER Martine, 2005, "Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German", *in Proceeding of Interspeech 2005*, Lisbon, Portugal, p. 2453-2456.

GENDROT Cédric, ADDA-DECKER Martine & VAISSIÈRE Jacqueline, 2008, « Les voyelles /i/ et /y/ du français : focalisation et variations formantiques », *in Actes des 27èmes Journées d'Études sur la Parole*, Avignon, p. 205-208.

KAMIYAMA Takeki, 2011, "Pronunciation of French vowels by Japanese speakers learning French as a foreign language: back and front rounded vowels /u y ø/", *Phonological Studies* (Phonological Society of Japan), vol. 14, p. 97-108.

KAMIYAMA Takeki, 2012, « Production des voyelles du français par des apprenants japonophones : effet du dialecte d'origine », *in* Laurent Besacier, Benjamin Lecouteux & Gilles Sérasset (dir.), *Actes des Journées d'Études sur la Parole 2012*, Grenoble, p. 771–778.

KAMIYAMA Takeki, KANEDA Jumpei & HAYASHI Ryôko, 2013, « Perception et production des voyelles orales du français par des japonophones avec des dialectes différents », *19ème Congrès International des linguistes / 19th International Congress of Linguists* (juillet 2013), Genève, Suisse.

KAMIYAMA Takeki & VAISSIÈRE Jacqueline, 2009, "Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners", *AILE … LIA*, n° 2, p. 9-41.

KOCJANČIČ ANTOLÍK Tanja, PILLOT-LOISEAU Claire & KAMIYAM Takeki, 2019, "The effectiveness of real-time ultrasound visual feedback on tongue movements in L2 pronunciation training", *Journal of Second Language Pronunciation*, vol. 5(1), p. 72-97.

KOKURITSU KOKUGO KENKYÛJO [National Institute for Japanese Language and Linguistics] (UEMURA Yukio & TAKADA Shôji), 1990, *Nihongo no boin, shiin, onsetsu: chouon undou no jikken-onseigakuteki kenkyû* [Vowels, consonants and syllables in Japanese: an experimental phonetic study on articulatory gestures], Tokyo, Shûei shuppan.

LABRUNE Laurence, 2012, *The Phonology of Japanese*, Oxford, Oxford University Press.

LADO Robert, 1964, *Language Teaching: A Scientific Approach*, New York, McGraw-Hill.

LANDRON Simon, PAILLEREAU Nikola, NAWAFLEH Ahmad, EXARE Christelle, ANDO Hirofumi & GAO Jiayin, 2010, "The PhoDiFLE Corpus: a Joint Corpus of French Speech, for a Phonetic Study of L1-L2 Contrasts", *Les cahiers de praxématique,* vol. 54-55, p. 73-86.

LIÉNARD Jean-Sylvain, 1977, *Les processus de la communication parlée : introduction à l'analyse et la synthèse de la parole* (Vols. 1-1), Paris ; New York ; Barcelone [etc.], Masson.

MAEDA Shinji, 1982, "A digital simulation method of the vocal tract system", *Speech Communication* vol. 1(3-4), p. 199-229.

MAEKAWA Kikuo, 2023, "Production of the utterance-final moraic nasal in Japanese: A real-time MRI study", *Journal of the International Phonetic Association*, n° 53(1), p. 189-212.

MARUSHIMA Naoki, DETEY Sylvain & KAWAGUCHI Yuiji, 2010, « Caractéristiques phonétiques des voyelles orales arrondies du français chez des apprenants japonophones », *Flambeau* (Revue annuelle de la section française, Université des Langues Étrangères de Tokyo) vol. 36, p 53–72.

MOKHTARI Parham & TANAKA Kazuyo, 2000, "A Corpus of Japanese Vowel Formant Patterns", *Bulletin of Electrotechnical Laboratory*, n° 64 (special issue), p. 57-66.

MORRISON Geoffrey Stewart, 2008, "Perception of synthetic vowels by monolingual Canadian-English, Mexican-Spanish, and Peninsular-Spanish listeners", *Canadian Acoustics*, n° 36(4), p. 17-23.

POLIVANOV Evgenij, 1931, « La perception des sons d'une langue étrangère », *Travaux du Cercle linguistique de Prague*, vol. 4, p. 79-96.

SCHWARTZ Jean-Luc, BOË Louis-Jean, VALLÉE Nathalie & ABRY Christian, 1997, "The Dispersion-Focalization Theory of vowel systems", *Journal of Phonetics,* vol. 25(3), p. 255-286.

SHIBATANI Masayoshi, 1990, *The Languages of Japan*, Cambridge [England], Cambridge University Press.

SUGITÔ Miyoko, 1995, *Oosaka - Toukyou akusento onsei jiten CD-ROM: kaisetsuhen* [CD-ROM Accent dictionary of Spoken Osaka and Tokyo Japanese], Tokyo, Maruzen.

TAKEBAYASHI Shigeru, 1996, *Eigo onseigaku* [English Phonetics], Tokyo, Kenkyûsha.

TRAUNMÜLLER Hartmut, 1990, "Analytical expressions for the tonotopic sensory scale", *Journal of the Acoustical Society of America,* vol. 88(1), p. 97-100.

TREMBLAY Marie-Claude & KAMIYAMA Takeki, 2009, "Perception of L2 production by L1 speakers of different dialectal backgrounds: the case of Japanese-speaking learners' /u/ perceived by French and Quebec native speakers. Cross-Language Speech Perception and Variations in Linguistic Experience" (2nd ASA Special Workshop on Speech, May 2009, Portland, Oregon, United States), *Journal of the Acoustical Society of America,* vol. 125(4), p. 2776.

TRUBETZKOY Nikolai S., [1939] 1969, *Principles of Phonology* (*Grundzüge der Phonologie,* translated by Christiane A. M. Baltaxe), Berkeley/Los Angeles, University of California Press.

TUBACH Jean-Pierre (dir.) (CALLIOPE), 1989, *La parole et son traitement automatique*, Paris, Masson.

VAISSIÈRE Jacqueline, 2006, *La phonétique*, Paris, Presses Universitaires de France.

VAISSIÈRE Jacqueline, 2007, "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", *in* Maria-Josep Solé, Patrice Speeter Beddor & Manjari Ohala (eds.), *Experimental Approaches to Phonology*, Oxford, Oxford University Press, p. 54-71.

VAISSIÈRE Jacqueline, 2011, "On the acoustic and perceptual characterization of reference vowels in a cross-language perspective", *in* Wai Sum Lee & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, City University of Hong Kong, p 52-59.

VANCE Timothy J., 2008, *The Sounds of Japanese,* Cambridge, Cambridge University Press.

WEINREICH Uriel, [1953] 1968, *Languages in Contact,* The Hague, Mouton.

Espaces Linguistiques

Thinking outside the guttural box: fresh perspectives on vowel lowering in Qeltu

L'abaissement vocalique en Qeltu : une approche sociophonétique

Abdulkareem YASEEN

University of Diyala

abdulkareem.env.hum@uodiyala.edu.iq

Abstract: This study revisits the process of vowel lowering in the Arabic dialect known as Qeltu, which is spoken in parts of northern Iraq and neighboring Syria and Turkey. It aims to uncover hitherto unreported patterns of vowel lowering and re-examines a frequently held axiom that attributes a categorical impact of specific linguistic units on the behavior of this phonological process (e.g. Al-Ani, 1970; Thelwall, 1990; Rose, 1996). The analysis of spoken data from 30 speakers considers various linguistic and social constraints that may or may not influence the process of vowel lowering. The data was collected through sociolinguistic interviews (Tagliamonte, 2006). It is generally acknowledged that the main allophones of Iraqi Arabic [i], [iː], [uː] (Al-Khalesi, 2006) are respectively realized as [a], [ɛː], and [ɔː] in Qeltu (Jastrow, 1994 & 2006). This is especially noticeable when these vowels are adjacent to guttural sounds, as exemplified in the words: /miʃhuːr/ *famous*, /sˤtˤuːħ/ *roof* and /daqiːq/ *flour*, which are pronounced as [maʃhɔːr], [sˤtˤɔːħ] and [daqɛːq], respectively (Jastrow, 1994; Muhammad, 2018). Results indicate that the relationship between guttural sounds and the lowering of vowels in Qeltu cannot be solely predicted by their proximity to each other, and that additional social factors are involved in determining this phenomenon.
Keywords: Qeltu, vowel lowering, Iraqi Arabic, sociophonetics, guttural consonants

Résumé : La présente étude traite du processus d'abaissement vocalique en Qeltu, une variété d'arabe, parlée dans le nord de l'Irak ainsi qu'en Syrie et en Turquie. Elle explore le fonctionnement de ce processus phonologique, visant à aller au-delà des explications communément associées à cette variété qui s'appuient exclusivement sur des considérations linguistiques (cf. e.g. Al-Ani, 1970 ; Thelwall, 1990 ; Rose, 1996). Les analyses se basent sur un corpus de parole spontanée de 30 locuteurs (selon la méthode des entretiens en sociolinguistique, Tagliamonte, 2006) et prennent en compte des facteurs tant linguistiques que socio-linguistiques susceptibles d'être corrélés à l'abaissement vocalique. Al-Khalesi (2006) indique que les allophones de l'arabe iraquien [i], [iː], [uː] sont réalisées respectivement [a], [ɛː], et [ɔː] en Qeltu (Jastrow, 1994 & 2006), surtout lorsque les voyelles sont précédées d'une consonne gutturale : /miʃhuːr/ *connu*, /sˤtˤuːħ/ *toit* et /daqiːq/ *farine*, seraient prononcés [maʃhɔːr], [sˤtˤɔːħ] et [daqɛːq] (Jastrow, 1994 ; Muhammad, 2018). Cependant, les résultats de cette étude montrent que la proximité avec une gutturale ne suffit pas à rendre compte de ce phénomène d'abaissement et que des facteurs sociolinguistiques jouent un rôle majeur.
Mots clés : Qeltu, abaissement vocalique, arabe iraqien, sociophonétique, consonnes gutturales

1. Vowel lowering: an overview

Vowel lowering is a process that involves a change in the height of vowel sounds within the vowel space. This change typically results in vowels being produced at a lower position in the space than they are otherwise. This phenomenon has been observed in a wide range of languages across the world, indicating its widespread occurrence and importance in phonology (Rose, 1996). English is one of the languages that display this phenomenon. Hickey (2018) describes varieties of English with lowering of short front vowels, i.e. *bet* for *bit*, *dack* for *deck* (Canadian, Californian and Dublin Englishes). Previous research (e.g. Labov, 1986; Hoffman, 2010; Eckert, 2012; Hickey, 2003) focusing on the English language has shown that the process of vowel shifting can have social or cultural implications beyond purely linguistic aspects. For instance, this phenomenon may be used by speakers to establish a sense of affiliation or dissociation with a particular group or community (Housholder, 1983; Rakić et al., 2011). This suggests that vowel lowering may not solely be a linguistic phenomenon, but also a sociolinguistic marker pointing at ways in which language is used in various social contexts.

Vowel lowering has also been observed in certain varieties of Arabic. However, Alshammari (2019) indicates that they have been predominantly studied either through a phonological lens (e.g. Rose, 1996) or a phonetic one (e.g Al-Ani, 1970, p. 34). Other studies have honed in on pharyngeal coarticulation and determined the patterns and mechanisms involved in vowel lowering in Arabic. These include Ali & Daniloff (1972), McCarthy (1994), Altairi *et al.* (2017) and Al-Ansari & Kulikov (2023). These studies, while considered essential contributions to the study of Arabic linguistics, are more theory-driven than data-driven. Indeed, the amount of linguistic data used in these studies is relatively limited. Nonetheless, these investigations represent an important starting point for further exploration and understanding of vowel lowering and other phonological features in Arabic.

2. The Qeltu variety

2.1. Geographic location

Qeltu is a dialect of Arabic spoken in a region commonly referred to as northern Mesopotamia. This area encompasses a vast expanse that spans northern Iraq and the parts around the conterminous borders with Syria and Turkey. In the current study, which focuses on the Iraqi part, Qeltu is found predominantly in towns situated along the Tigris and Euphrates rivers. Accordingly, in Jastrow's (2006b) trichotomous classification of the Qeltu dialects, three distinct subgroups are identified: Tigris, Euphrates, and Anatolian. The Euphrates group of Qeltu dialects extends along

the Euphrates basin, encompassing towns situated along the Euphrates River, including locations like Hit and Ana (as illustrated[42] in Figure 1).

Figure 1: Qeltu-speaking area (North Mesopotamia)
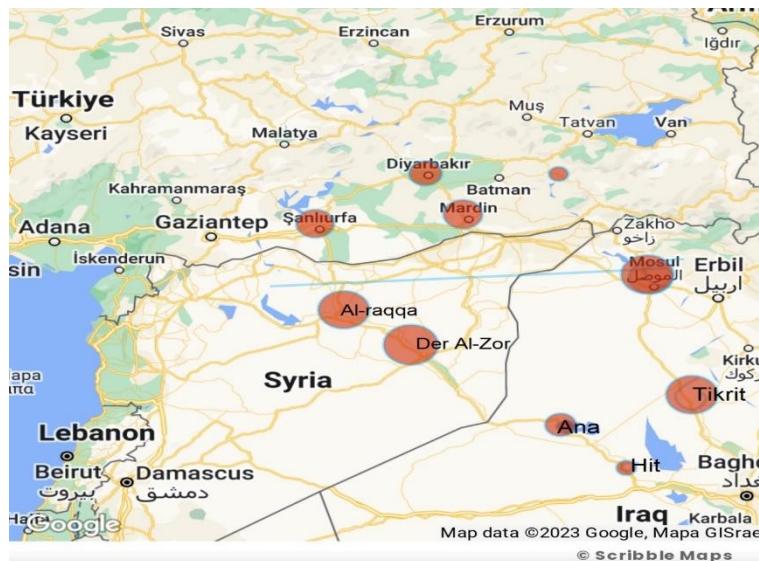
2.2. The two Qeltu varieties analysed in this study

The aim of this paper is twofold: to present new findings on vowel lowering in two urban Qeltu dialects using the author's dataset to shed light on the variation within this process. The first dialect is spoken in the city of Mosul, Iraq, Iraq's third largest city. Regardless of their religion or ethnicity, all participants of the study are native Maṣlāwī (i.e. native of Mosul) and speak Qeltu. The use of Qeltu in Mosul and Baghdad differs in that Qeltu is spoken according to religious affiliation in Baghdad, with Jews and Christians using it while Muslims speak Gilit (Blanc, 1964). The Qeltu dialect spoken in Mosul belongs to the Tigris group of Qeltu dialects. This classification is based on the fact that Mosul is one of the towns located along the Tigris river. The other Qeltu dialect considered in this study is Hiti Arabic (henceforth HA). This dialect is spoken in the town of Hit, which is situated around 180 km to the west of Baghdad on the Euphrates river. The dialect of Hit belongs to the Euphrates group of Qeltu dialects.

2.3. Phonetic specificity of Qeltu: vowel lowering

Qeltu-speaking communities tend to lower their articulation of /uː/ to [oː] as in /suːq/ → [soːq] *'market'*. Vowel lowering is a well-documented aspect of the Anatolian Qeltu Arabic dialects, which

---

[42] The maps were created by the author using Google Maps and Microsoft paint software as an editor.

are spoken in the Anatolian region of Turkey. A number of studies have reported this phenomenon–Lahdo (2009) focuses on the Tillo dialect spoken in the Siirt province of Turkey. The fundamental rule governing vowel lowering is that when an emphatic consonant or any of the consonants /x/, /ɣ/, /q/, /ħ/, /ʕ/ is adjacent to the vowels /iː/ and /uː/, the vowels undergo a process of lowering, resulting in their pronunciation as [eː] and [oː] respectively. This can be illustrated in the example of the words /maluːħ/ (*good-looking*) in which the vowel is lowered resulting in [maloːħ]. Jastrow (2006a) identified a similar pattern of vowel lowering in the Anatolian Arabic of Kinderib where the word for *flour* is pronounced as [daqeːq] instead of the standard /daqiːq/. Jastrow attributes the occurrence of lowering in the dialect to the influence of emphatic consonants.

The spoken form of Qeltu in Iraq has also been found to exhibit vowel lowering. In Iraqi Arabic, Erwin (1963, p. 24) indicates that the pronunciation of vowels such as /uː/ is influenced by the presence of neighbouring emphatic sounds. When the vowel comes in a context contiguous to any of these sounds, it is usually realized as a back, close, and long vowel, with a small variation in pronunciation. Conversely, when the vowel is situated near non-emphatic sounds, it may be pronounced differently. While lowering is a defining feature of Qeltu, Blanc (1964, p. 41) notes that he could not find it in certain verbs such as [asuːq] and [abuːq], which mean *'I drive'* and *'I steal'*, respectively. Likewise, the aggregate of studies only briefly reported this phenomenon in Qeltu without assessing it with a thorough analysis. Oussani (1901, p. 101) made the first mention of this phenomenon citing it in the speech of Baghdadi Christians as an outcome induced by, albeit rather tenuously, contiguous guttural consonants. This statement gained traction in subsequent accounts (e.g. Blanc, 1964; Jastrow, 1994, 2006b) who have observed that /uː/ and /iː/ are realized as [oː] and [eː] correspondingly in guttural contexts, as in the words [xjoːt] *threads* and [daqeːq] *flour* (Jastrow, 2006a, p. 417), as against their standard Gilit forms of [xjoːt] and [daqiːq] respectively. Abu Haidar (1991, p. 18) notes that lowered articulations are more frequently produced in Christian Baghdadi Qeltu than in the Jewish Baghdadi one. Talay (2011, p. 913) reports this type of lowering in the speech of Jews in the northern part of Iraq. The purpose of this study is to further investigate this phenomenon considering two different Qeltu varieties of Arabic.

3. Dataset & Method

3.1 Selection of informants and recording protocol

The Mosul data comes from recordings of informal conversations of 30 participants. Each age cohort had five speakers, distributed evenly across gender categories. The entire sample consisted of 15 individuals from the Middle Class and 15 from the Lower Middle Class. Establishing a reliable social classification proved to be both theoretically and methodologically challenging due to the scarcity of published information on social classifications in Mosul, compounded by the evolving societal landscape. However, as Watt (1998, p. 98) highlights, the knowledge possessed by insiders of the community, whether they are ordinary residents or researchers, can serve as a highly dependable indicator of existing class distinctions within the community. Therefore, the participants recruited for this study were classified based on their respective neighbourhoods, as these emerged as the most reliable indicators for categorizing Mosul residents in terms of their social class. This approach was further informed by insight provided by the informants who possess intimate knowledge of these areas. As a result, each speaker's categorization was confirmed by the speakers themselves, who self-identified as either lower middle-class or middle-class according to the neighbourhood in which they lived. The neighborhoods on the left coast (Figure 2) were identified as middle class while those on the right coast were identified as lower middle class.

Figure 2: Qeltu-speaking areas inside Mosul (on both the Right & Left coasts)
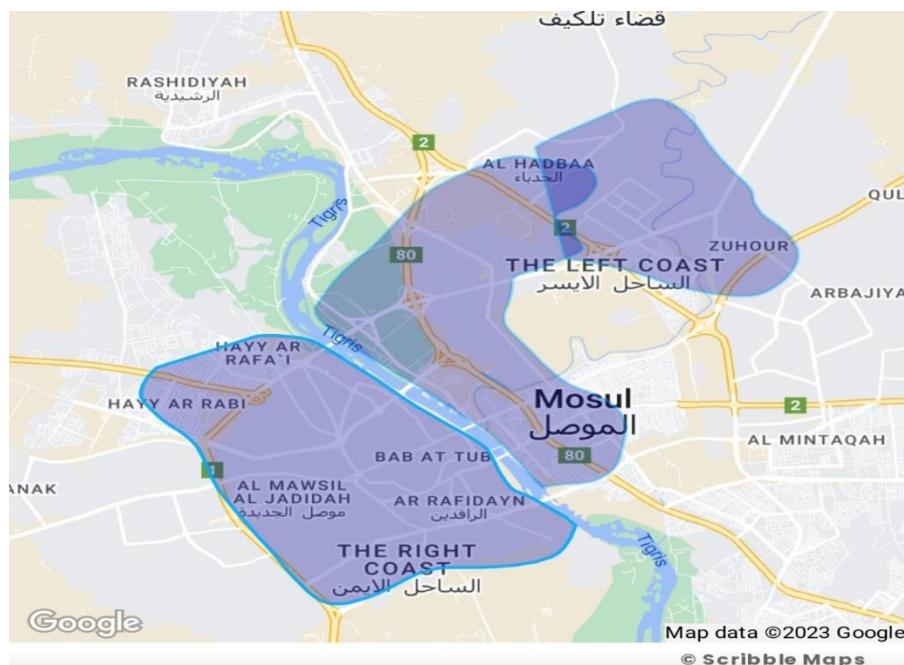
The recordings for this study were conducted within the participants' residences. In order to create a noise-free environment, small rooms furnished with items such as curtains to minimize acoustic

reverberations were used for the recording sessions. Digital recording was used with a sampling rate of 44.1 kHz and a 16-bit resolution, using a portable Edirol R-09HR High-Resolution recorder and an Edirol CS-50 condenser microphone placed on a desk in front of the participants. Before engaging in the study, every participant signed informed consent, confirming their understanding of the scope of their participation. The duration of the sessions varied among speakers, although it was ensured that each session lasted approximately one hour or slightly longer, provided that the speakers were comfortable with the interview and the topics under discussion. Following the approach outlined by Tagliamonte (2006, p. 39), the questions were adapted to include the informants' preferences in conversation topics, thus enabling a smooth transition between a variety of topics ranging from general questions to more personal themes. This contributed to creating a controlled but accommodating environment for the participants so as to limit the observer effect (Labov, 1981).

3.2. Targeted tokens

Around 30 hours of natural speech comprised the dataset, from which tokens were extracted and coded for subsequent analysis. In line with Wolfram (1993, p. 214), the number of tokens has been limited to a maximum of 3 per participant to ensure that the dataset includes a wide range of lexical elements and avoids repetitive or overly common items. As such, a total of (1995) lexical items were included in the analysis. Table 1 below gives a sample of tokens coded for analysis.

Table 1: Sample of words retrieved for analysis

Word	*qeltu*	*gilit*	Gloss
quul	[qu:l]	[gu:l]	say
qasseer	[qasˤɪːɣ]	[gisˤɪːr]	short
qaed	[qɛːʕid]	[gaːʕid]	sitting
lahaqtu	[laħħaqtu]	[laħħagit]	I followed
urooq	[ʔɪʕɣɔːq]	[ʔɪʕruːg]	a type of bread
qroon	[qruːn]	[gruːn]	horns
maslooq	[maslɔːq]	[masluːg]	boiled
fooq	[fɔːq]	[fɔːg]	up
sooq	[sɔːq]	[suːg]	market
qasaseeb	[qasˤɛːsˤiːb]	[gisˤaːsˤiːb]	butchers
sandooq	[sˤanduːq]	[sˤanduːg]	box
yethooq	[ʔɪðuːq]	[ʔɪðuːg]	tasting
aqraba	[ʕaqqabi]	[ʕagruba]	scorpion
yeaalqoon	[yiʕalquːn]	[yiʕalguːn]	hanging
tqoom	[tqɔːm]	[tguːm]	standing up
yetbaq	[yitˤbuq]	[yitˤubg]	standing by
qumtu	[qumtu]	[gumit]	I did

3.3 Data Analysis

Each lexical item was coded according to the predictors selected in this study. The coding of the tokens was done taking into consideration approaches found in previous studies and the objectives of the current study. The tokens were coded for three factors: the part of speech, the preceding sounds and the following sounds. The aim of this assessment was to determine the extent to which the vowel was lowered. The values of the first two formants were measured precisely at the midpoint of each token. This was done to avoid potential influence from adjacent segments (Klatt and Klatt, 1990, p. 829; Thomas, 2010, p. 148). The acoustic measures were extracted through the open-source software 'Praat' (Boersma & Weenink, 2017). A Praat script was used to segment and annotate all the coded tokens (García, 2017). During this procedure, both sound and visual cues were used to help determine the beginning and end points of each targeted vowel segment, taking into account the guidelines outlined by Thomas (2010) regarding where to place the boundaries between the adjacent sounds that surround the vowel. After completing the segmentation, all tokens were manually inspected. Approximately 30% of the segmented data underwent a secondary review conducted by a second rater who is trained in phonetics and speaks Arabic. The rater validated 96% of the measurements and suggested alterations for the remaining 4%. Subsequently, a meeting was convened with both raters to reach a consensus on the positioning of boundaries. Labov's Telsur method was used to normalise the vowel formant data, and this process was carried out using the online resource NORM Suite (Thomas *et al.*, 2007). Labov's Telsur method is a revised version of Nearey's method and has proven to be more effective than other methods tested in terms of producing accurate data without distortions and preserving the sociolinguistic variation. Telsur uses the log mean to normalize the formant values, which allows for consistent comparisons across speakers, and then calculates a single grand mean to represent the speech characteristics of the entire group under study. This method has been shown to be highly reliable and useful in sociolinguistic research. A statistical analysis was then carried out to determine the frequency and distribution of the responses (F1 & F2) across the various parameters (i.e. social and linguistic predictors) that were hypothesized to have an impact on them. The goal was to assess the interplay between the parameters and the responses and determine the significance of any observed patterns.

4. Results

4.1. Results of vowel lowering in Mosul

The data was analysed using Rbrul[43] (Johnson, 2009). A mixed-effect step-wise regression analysis was carried out to evaluate the impact of different factors (social and phonetic) on vowel lowering. A model was designed in which all the contexts discussed above were included as fixed predictors. In this model, speaker and lexical item were retained as random effects. Formant values (F1 then F2) serve as the dependent variable. The Rbrul runs conducted for the first formant are displayed in Table 2 below. The table contains a list of the significant and non-significant predictors that emerged from the step-down model. Related values are listed for each predictor. These include the linear coefficients, the number of lexical items coded and the F1 mean value. The data presented in the table reveals that only age emerged as a significant factor impacting vowel height. Indeed, younger and middle-aged speakers exhibit a tendency towards using variants of this vowel that are higher in quality when compared to the group of older speakers. This is evidenced by the contrasting negative coefficients and mean values observed in these two groups, as opposed to the older group, the coefficient of which was positive and the mean formant value was greater. Evidently, this underscores the dynamic shift in the height of this vowel over time.

It is also evident from the results that the height of /uː/ is not significantly conditioned by the sound that follows or precedes it. This shows that there is no favourable context (e.g. guttural) for the lowered rendition of /uː/ as has been suggested by previous accounts.

---

[43] Rbrul, also known by the name "brul," functions as an R package utilized for the analysis of linguistic data, with a particular focus on the disciplines of linguistics and sociolinguistics. Its primary design is aimed at streamlining the examination of language variation and the dynamics of language evolution through the utilization of mixed-effect modelling techniques. Researchers employ Rbrul to investigate linguistic phenomena, including but not limited to phonological or syntactic shifts, while also scrutinizing the impact of various factors on patterns of language usage. Johnson (2009) reasons that the prevailing edition, GoldVarb, which is widely used, exhibits limited adaptability and, furthermore, separates its users from the broader community of quantitative linguists. Rbrul, being a fresher version of the variable rule software, seeks to tackle these issues.

Table 2: Rbrul results for F1 in /uː/. Mixed-effect effect model with F1 (Hz) as y-variable and linguistic factors as x-variable. Random effects: speaker & lexical item

Factor		Coef	Tokens	Mean
Age	Old	50.464	621	509.424
	Middle	-1.750	699	457.101
	Young	-48.714	675	407.776
Speaker mean	456.699			
Intercept	459.266			
Df	5			
Tokens Total	1995			

Non-significant factors (dropped from best model)

Factor		Coef	Tokens	Mean
Gender	Female	[N.S.]	964	474.121
	Male	[N.S.]	1031	464.218
Class	LMC	[N.S.]	1157	474.032
	MC	[N.S.]	838	462.06
Part of Speech	Adjective	[N.S.]	138	473.45
	Noun	[N.S.]	462	470.276
	Verb	[N.S.]	1395	468.141
Preceding Sound	Affricate	[N.S.]	76	477.994
	Approximant	[N.S.]	411	467.819
	Emphatic	[N.S.]	142	474.261
	Fricative	[N.S.]	251	467.528
	Nasal	[N.S.]	276	465.278
	Pharyngeal	[N.S.]	141	477.186
	Stop	[N.S.]	427	469.979
	Uvular	[N.S.]	271	464.887
Following Sound	Approximant	[N.S.]	193	475.033
	Emphatic	[N.S.]	114	459.675
	Fricative	[N.S.]	294	471.383
	Nasal	[N.S.]	980	469.162
	Pharyngeal	[N.S.]	164	462.952
	Stop	[N.S.]	96	458.479
	Uvular	[N.S.]	154	475.801

The boxplots in Figure 3 demonstrate the variability of F1 in relation to the phonological context. The height of the vowel /uː/ does not seem to be determined exclusively by a particular context, as there are no discernible differences between the spread of boxplots. Despite the fact that this vowel's realization is said to be lower next to pharyngeal contexts, this was not found to be statistically significant. The boxplots of the contexts coded in the model seem to be fairly comparable, with no marked differences between them.

Figure 3: Boxplots of F1 (Hz) in /u:/ according to the predictor of following sound
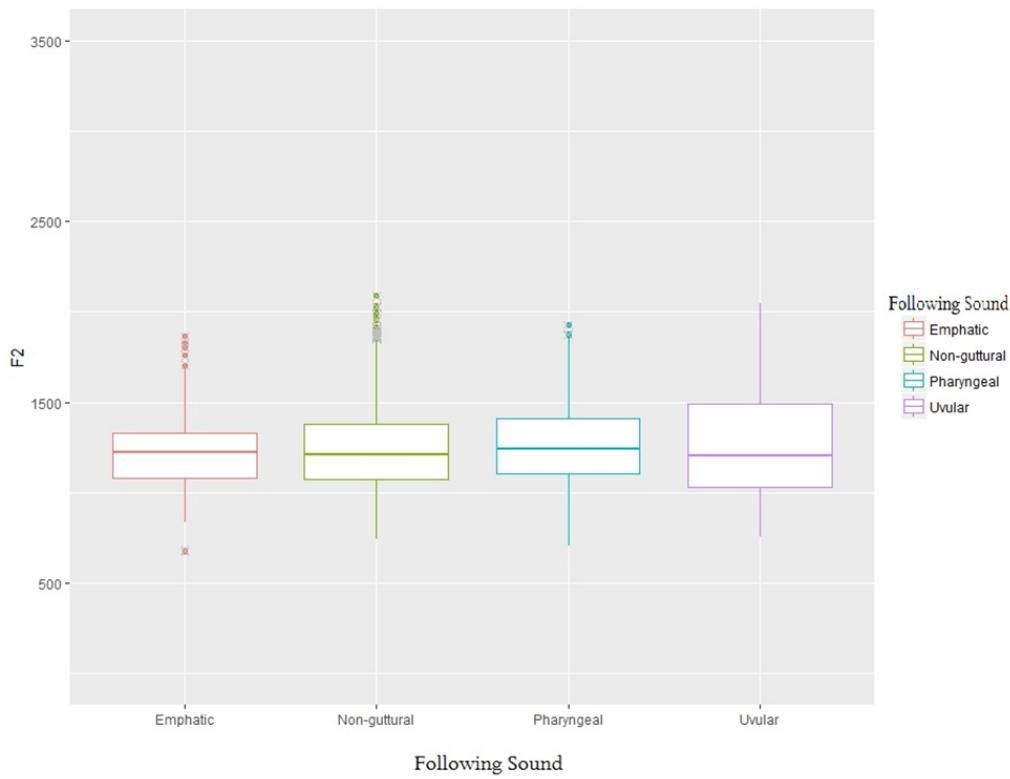

Table 3 displays the Rbrul runs for F2 as a dependent variable. It illustrates the advancement dimension of this vowel which suggests that there is a change in apparent time (i.e. across age-groups) as indicated by a statistically significant difference observed between the younger cohort and the middle-aged and elderly groups. The younger speakers, in particular, demonstrate a tendency towards a more back quality of the vowel (lower F2 compared to the other age groups), as indicated by a negative coefficient (-58.347), in contrast to the positive coefficients observed among the middle-aged and elderly groups.

The table also reveals that the vowel is fronted when preceded by guttural sounds (pharyngeals, uvulars, and emphatics). This stands in contrast to its pronunciation adjacent to non-guttural sounds. However, it is unexpected to encounter a rise in the F2 value, given the retraction of the tongue root and dorsum during the production of these sounds which commonly influences the surrounding vowels (Shahin, 2002). Moreover, the impact of these sounds on adjacent vowels depends on both the type of vowel and the particular dialect under investigation. When the vowel /u:/ is articulated adjacent to a guttural sound, such as a pharyngeal, the tongue's posture shifts downward and forward, leading to noticeable increases in F1 and F2 values. Consequently, this vowel tends to be pronounced as lowered and centralised (Vaissière, 2011, p. 10). The results of this study align with a number of previous investigations. Al-Ani (1976), in his examination of the

influence of pharyngeal sounds on the formants of neighbouring vowels in Iraqi Arabic, assessed the pharyngeal sound /ʕ/ and noted its tendency to raise the second formant of the /u/ vowel while lowering it for /a/. Similar observations have also been documented by Butcher and Ahmad (1987) in the context of Iraqi Arabic and by Alwan (1986) in various Arabic dialects, including Iraqi Arabic.

This research has also demonstrated that the fronted variant of this vowel is more often pronounced in verbs and adjectives compared to nouns. Although there is a scarcity of prior investigations available for comparative analysis, the sole plausible explanation for this phenomenon might be associated with it being an artefact of the dataset of this study. Consequently, any generalization must be approached with caution. Further research is imperative to elucidate the underlying nature of this observed pattern. Moreover, any form of labialisation will result in a general lowering of formant values (cf. Fant, 1960 and illustrated in Vaissière, 2009).

Table 3: Rbrul results for F1 of /uː/. Mixed-effect model with F1 (Hz) as y-variable and social & linguistic factors as x-variable. Random effects: speaker & lexical item

Factor		Coef	Tokens	Mean
Age	Middle	55.505	699	1261.250
	Old	2.842	621	1215.545
	Young	-58.347	675	1158.806
Part of speech	Verb	31.362	1395	1221.522
	Adjective	2.930	38	1205.690
	Noun	-34.292	462	1186.697
Following Sound	Uvular	34.850	154	1234.797
	Pharyngeal	5.956	164	1243.967
	Emphatic	2.856	114	1195.856
	Non-Guttural	-43.662	1563	1208.039
Speaker mean	1212.362			
Intercept	1210.717			
Df	5			
Tokens Total	1995			

Non-significant factors (dropped from best model)				
Gender	Female	[N.S.]	964	1246.307
	Male	[N.S.]	1031	1245.821
Class	LMC	[N.S.]	1157	1233.922
	MC	[N.S.]	838	1262.808
Preceding Sound	Affricate	[N.S.]	76	1189.682
	Approximant	[N.S.]	411	1225.968
	Emphatic	[N.S.]	142	1226.277
	Fricative	[N.S.]	251	1275.925
	Nasal	[N.S.]	276	1217.916
	Pharyngeal	[N.S.]	141	1232.607
	Stop	[N.S.]	427	1262.784
	Uvular	[N.S.]	271	1284.331

Examining the combined F1 and F2 patterns reveals a distinct preference among young speakers for a vowel realization characterized by increased height and retraction, akin to a Gilit-like quality. Figure 4 below illustrates this comprehensive shift in vowel articulation, showcasing both F1 and F2 measurements to map out the articulatory patterns within the vocalic space. A noticeable tendency among young speakers is evident, as their realizations tend to cluster towards a higher and more retracted position on the plot. Conversely, the plot highlights that the articulations of the middle-aged cohorts tend to cluster at a lower position compared to the younger group. Notably, there is a more pronounced distinction between the young and elderly groups, with the latter's productions occupying a lower position on the plot.

Figure 4: A vowel plot illustrating the behaviour of /uː/ according to the age groups incorporated in the study

Thus, what appears to be at stake in the behaviour of this vowel is its sociolinguistic dimension, with the linguistic aspect showing no significant impact, as indicated by statistical measurements.

As regards /iː/ lowering, through the analysis of the data, it was found that this vowel was produced as /iː/ rather than the often-reported lowered form (i.e. /eː/) almost systematically. Only one token was found to be realized with the lowered version of this vowel [ʕaqeːq] *'agate'*. This was found in the speech of a +60 male speaker. This almost systematic patterning found amongst the participants suggests that this phenomenon has undergone change and is nearing completion.

Moreover, a further exploration of the data revealed that lowering of /uː/ is observed in a wider range of verb types than previously noted in the literature (Blanc, 1964). As such, its occurrence cannot be restricted to the verb conjugations reported by Blanc (1964, p. 41) i.e. [asuːq] *'I drive'* and [abuːq] *'I steal'*, traditionally pronounced [asoːq] and [aboːq] in Qeltu.

It can also be noted that lowered articulation of /uː/, i.e. [oː] can actually occur in contexts other than those of gutturals. This also stands in contrast to previous accounts (e.g. Jastrow, 1994; Oussani, 1901). In fact, the data revealed that lowered productions of /uː/ can occur across the board in terms of neighbouring segments even in entirely guttural-free tokens as can be illustrated in Table 4:

Table 4: Examples of /uː/ lowering in different phonological contexts

Word	Preceding sound	Following sound	Gloss
[tfoːʁ]	[+labiodental]	[+uvular]	'it boils'
[yəbkoːn]	[+velar]	[+alveolar]	'they cry'
[tmoːt]	[+bilabial]	[+alveolar]	'she dies'
[falloːhaˑ]	[+alveolar]	[+glottal]	'they unfastened it'

In view of the results discussed above, the case of Mosul contributes to the discussion in a number of ways, as follows:

1) Re-evaluation of previous accounts: the findings contradict earlier assumptions that proposed specific conditions governing vowel lowering. The results demonstrate that the height of /uː/ is not necessarily influenced by certain adjacent sounds.

2) Highlighting systematic patterns: the nearly systematic pronunciation of /iː/ as [iː], with just one exception, points to a potential, completed linguistic change, which is significant for phonological, sociophonetic and sociolinguistic research in Arabic.

3) Expanding the scope of contextual understanding: the results reveal that the phenomenon of /uː/ lowering extends beyond specific verb conjugations or guttural contexts and occurs across various verb types.

4.2. Results of vowel lowering in Hit

For the sake of comparison with other Qeltu dialects, I present findings from a study by Mohammed (2018) in Hiti Arabic (HA) that includes vowel lowering. Mohammed discusses this

phenomenon noting that it is generally considered as a coarticulation process where a front vowel /i/ is lowered in the vicinity of emphatic and guttural consonants as in [qisˤsˤaːb] vs [qasˤsˤaːb] *butcher*. However, his analysis produced findings that are square with those reported for Mosul above. Results showed that while HA displays vowel lowering, it is nevertheless not solely restricted to emphatic or pharyngeal environments. Table 5 below displays words with vowel lowering (in italics) occurring next to sounds other than guttural ones.

Table 5: Words exhibiting vowel lowering of /i/ in Hiti Arabic

qiltu (HIA)	*gilit*	English gloss
[ti̇nnuːr]	[tɑnnuːr]	Oven
[mis̥ˤnuːʕ]	[mɑsˤnuːʕ]	Produced
[miʃhuːr]	[mɑʃhuːr]	Famous
[nidʒdʒaːr]	[nɑdʒdʒaːr]	Carpenter
[hiddaːdiːn]	[hɑddaːdiːn]	Ironmongers
[fillaːh]	[fɑllaːh]	Farmer

Source : Mohammed, 2018, p. 138

In his analysis of /i/ lowering to [a] in unstressed closed syllables in the three age groups, Mohammed (2018) has demonstrated that this process occurs regardless of the adjacent consonants. He points out that his findings are surprising as the process of vowel lowering is often associated with back consonant environments (e.g. velars and emphatics). His findings, therefore, stand in contrast to earlier research on vowel lowering as they reveal that the presence of emphatics in a context with a high vowel does not inevitably result in an increase in the lowering of this vowel, but rather a decrease. The occurrence of vowel lowering in pharyngeal and non-pharyngeal environments is shown in Figure 5 below. The 'other consonants' group in the illustration here represents the following consonants: (/ʕ/, /ħ/ /χ/, /ɣ/ & /q/).

Figure 5: vowel lowering in pharyngeal and non-pharyngeal contexts
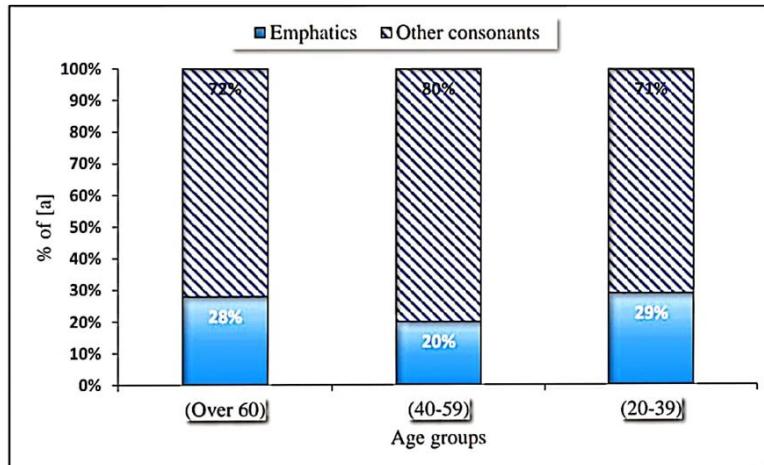
Source: Mohammed, 2018, p. 162

A noteworthy pattern can therefore be found in HA where Mohammed (2018) analyzed the phenomenon of the vowel /i/ being lowered to [a] in unstressed closed syllables within three age groups. Irrespective of the consonants preceding or following the vowel, a slight change in apparent time (i.e. categorized by age-group) is noticeable but lacks statistical significance (Table 6), in contrast to the findings for Mosul Qeltu.

Table 6: Vowel lowering in HA distributed according to age

Age Group	[a]	[i]	Subtotal
Over 60	28 % (101 tokens)	72 % (242 tokens)	343
40-59	39 % (180 tokens)	61 % (198 tokens)	378
20-39	41 % (150 tokens)	59 % (240tokens)	390
	ANOVA: F= 1.536; P= 0.230		Total: 1,111

Source: Mohammed, 2018, p. 160

An interesting aspect observable in vowel lowering in Hit is that this process of lowering is the supralocal (i.e *Gilit*) form towards which Hitis shift their speech while the reverse is true in Mosul where it represents the traditional form. Thus, while vowel lowering appears to be behaving in much the same way *vis-à-vis* its linguistic conditioning, the sociolinguistic patterning of this phenomenon is noticeably distinct. This implies that the phonetic environment has less influence than previously thought on vowel lowering and that sociolinguistic factors play a significant role.

These findings show how important it is to continue research on Qeltu varieties, substantial areas of which remain uncharted territory.

5. Discussion

Taking into consideration the variability patterns in this phonological process, this study has attempted to investigate some claims made in previous accounts. As we have seen earlier, both Qeltu dialects covered in this study display patterns that run counter to what has long been claimed in the literature (e.g. Oussani, 1901; Blanc, 1964; Jastrow, 1994). The findings of the current study testify to a crucial insight into the behaviour of the lowering of /i/, /u:/ and /i:/ in that there are other potential predictors besides the presence of guttural consonants whether before or after the vowel. Also, the findings of this study could not lend support to the lowering of /u:/ to [o:] in certain verbs as it appears that it is not limited to certain verbs, i.e. [asu:q] *'I convey'* and [abu:q] *'I steal'*. The results discussed in the current study have also demonstrated that lowered forms of /u:/ and /i/ can come in different lexical items and phonological contexts. Taken together, these results provide a rather different picture of what appeared to be an initially *prima facie* case of guttural-induced process of vowel lowering, particularly in Qeltu. The guttural-based proposal has been a perennial theme in the literature for a significant period of time. Although earlier studies did not elaborate on why the guttural sounds occupy the foreground of their argument, it must be said that the orientation and scope of interest are relative to the type of each study and the era in which it is carried out. Al-Wer (2014, p. 401) observes that, parallel to those early works in dialectology and sociolinguistics in Europe, Arabic also had its own limitations when it comes to methodology. The bulk of the previous research was heavily dependent on impressionistic data collected from a restricted pool of participants. Previous accounts aimed at simply capturing a general description of the linguistic profile of the dialects concerned.

As there is no reliable historical record of this phenomenon, it is difficult to assess how far vowel lowering in Qeltu is the result of gradual sound change. Generally, vowel lowering is a complex phenomenon with various underlying factors that impact its occurrence and patterns of use. Ohala (1981) argued that vowel lowering can be influenced by different factors, such as the sounds around the vowel, the social situation of the speaker, and the physical limitations of the vocal tract. While physiological constraints, such as the size and shape of the vocal tract, can contribute to vowel lowering, it is also undeniable that social factors such as age or gender can play a role in the patterning of this phenomenon.

The findings of Mosul and Hit emphasise the intricate nature of linguistic phenomena, particularly focusing on vowel lowering. These findings underscore the need for a more nuanced understanding of phonological processes so as to incorporate a myriad of factors. What needs to be also factored in is considering its behaviour in other Qeltu dialects, taking into account the availability of more rigorous analytical methods. This makes it possible to gain a deeper understanding of the underlying causes and evolution of vowel lowering in Qeltu, specifically with regard to /i/, /uː/, and /iː/ vowels.

Conclusion

Neither the current study nor previous accounts are the finished article on the subject of vowel lowering in Qeltu. A potential line for further investigation could be to study the language from a sociolinguistic perspective taking into account the effects of the long history of the region on the development of vowel lowering on Qeltu in particular and Iraqi Arabic in general. Our findings invite us to critically reassess previously accepted claims and assumptions and highlight the importance of multiple sociolinguistic factors and their significant influence on linguistic patterns. Age, for instance, appears to be playing a substantial role in shaping the behaviour of Qeltu speech. More precisely, the F1 and F2 patterns unveil a clear tendency among younger speakers for producing the vowels assessed in this study with a quality that is higher and more retracted, resembling a Gilit-like quality. They suggest that language is not isolated but deeply intertwined with social context, requiring a sociolinguistically informed perspective for a more comprehensive understanding. The studies of Mosul and Hit also shed light on regional linguistic variation, demonstrating that linguistic phenomena may exhibit diverse variation patterns across different dialects. Consequently, it is evident that what holds true in one dialect or region may not apply universally, thereby showing how important it is to consider dialectal differences in linguistic research.

References

ABU HAIDAR Farida, 1991, *Christian Arabic of Baghdad*, Wiesbaden, Harrassowitz.

AL-ANI Salman, 1976, *An acoustical and physiological investigation of the Arabic ʿayn*, Paris, Mouton.

AL-ANI Salman H., 1970, *Arabic phonology: An acoustical and physiological Investigation*, The Hague, Mouton.

AL-ANSARI Noora & KULIKOV Vladimir, 2023, "Acoustics of Arabic Uvulars and Emphatic Coronals: Evidence for Uvularization of Emphatics in Qatari Arabic", *Al-ʿArabiyya: Journal of the American Association of Teachers of Arabic*, vol. 55, p. 37-66.

AL-WER Enam, 2014, "Language and gender in the Middle East and North Africa", *in* Susan Ehrlich, Miriam Meyerhoff, & Janet Holmes (eds.), *The handbook of language, gender and sexuality*, Oxford, Wiley-Blackwell, p. 396-411.

ALI Latif Hassan & DANILOFF, Raymond G., 1972, "A contrastive cinefluorographic investigation of the articulation of emphatic–non-emphatic cognate consonants", *Studia Linguistica*, vol. 51, n° 2, p. 61-124.

ALSHAMMARI Mashaan, 2019, "Vowel lowering in Arabic: A phonological and phonetic study", *Journal of Language and Linguistic Studies*, vol. 15(4), p. 1133-1156.

ALTAIRI Hamad, BROWN Jason, WATSON Catherine, GICK Bryan, 2017, "Tongue Retraction in Arabic: An Ultrasound Study", *Proceedings of the 2016 Annual Meeting on Phonology*.

ALWAN Abeer A.-H., 1986, *Acoustic and perceptual correlates of uvular and pharyngeal consonants*, Unpublished thesis, Master's, Massachusetts Institute of Technology.

BLANC Haim, 1964, *Communal dialects in Baghdad*, Cambridge, Harvard University Press.

BOERSMA Paul & WEENINK David, 2017, *Praat: Doing phonetics by computer*, Amsterdam, Institute of Phonetic Sciences, University of Amsterdam.

BUTCHER Andrew & AHMAD Kusay, 1987, "Some acoustic and aerodynamic characteristics of pharyngeal consonants in Iraqi Arabic", *Phonetica*, vol. 44(3), p. 156-172.

ECKERT Penelope, 2012, "Northern California Vowels", Online at <https://web.stanford.edu/~eckert/vowels.html>

ERWIN Wallace M., 1963, *A short reference grammar of Iraqi Arabic*, Washington, DC, Georgetown University Press.

FANT Gunnar, 1960, *Acoustic theory of speech production*, the Hague, the Netherlands, Mouton & Co.

GARCIA Wendy, 2017, "Vowel Formants v1" [Praat script]. Retrieved from https://github.com/wendyelviragarcia/vowels

HICKEY Raymond, 2018, "'Yes, that's the best': Short front vowel lowering in English today: Young people across the anglophone world are changing their pronunciation of vowels according to a change which started in North America", *English Today*, vol. 34(2), p. 9-16.

HICKEY Raymond, 2003, "What's cool in Irish English? Linguistic change in Contemporary Ireland", *in* Hildegard L. C. Tristram (ed.), *Celtic Englishes III*, Heidelberg, Winter, p. 357-373.

HOFFMAN Michol F., 2010, "The role of social factors in the Canadian Vowel Shift: Evidence from Toronto", *American Speech*, vol. 85(2), p. 121-140.

HOUSEHOLDER Fred W., 1983, "Kyriolexia and language change", *Language*, vol. 59(1), p. 1-17.

JANSEN Sandra, 2018, "Peripheral Communities and Innovation: Changes in the GOOSE Vowel in a West Cumbrian Town", *in* Natalie Braber & Sandra Jansen (eds.), *Sociolinguistics in England*, London, Palgrave Macmillan, p. 297-331.

JASTROW Otto, 1994, "The Qeltu Arabic dialects of Mesopotamian Arabic", *Actas del Congreso Internacional sobre interferencias lingüísticas arabo-romances y paralelos extra-iberos: celebradas en Madrid*, Madrid. 10-14 December 1990, p. 119-124.

JASTROW Otto, 2006a, "Anatolian Arabic", *in* Kees Versteegh, Mushira Eid, Alaa Elgibali, Manfred Woidich & Andrzej Zaborski (eds.), *Encyclopedia of Arabic language and linguistics*, vol. 1, p. 87-96.

JASTROW Otto, 2006b, "Iraq", *in* Kees Versteegh, Mushira Eid, Alaa Elgibali, Manfred Woidich & Andrzej Zaborski (eds.), *Encyclopedia of Arabic language and linguistics*, vol. 2, Leiden: Brill, p. 414-427.

JOHNSON Daniel Ezra, 2009, "Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis", *Language and linguistics compass*, vol. 3(1), p. 359-383.

KLATT Denis H. & KLATT Laura C., 1990, "Analysis, synthesis and perception of voice quality variations among Female and male talkers", *The Journal of the Acoustical Society of America*, vol. 87 (2), p. 820-857.

LABOV William, 1981, "Field Methods of the Project on Linguistic Change and Variation", *Sociolinguistic Working Paper Number 81*.

LABOV William, 1986, "The social origins of sound change", *in* Harold B. Allen & Michael D. Linn (eds.), *Dialect and language variation*, Academic Press, p. 524-541.

LABOV William, 2001, *Principles of linguistic change. Volume 2: Social factors*, Malden, Massachusetts, Blackwell.

LAHDO Ablahad, 2009, *The Arabic Dialect of Tillo in the Region of Siirt, south-eastern Turkey*, Acta Universitatis Upsaliensis.

McCARTHY John J., 1994, "The phonetics and phonology of Semitic pharyngeals", *in* Patricia A. Keating (ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, p. 191-233.

MESTHRIE Rajend, 2010, "Sociophonetics and social change: Deracialisation of the GOOSE vowel in South African English 1", *Journal of sociolinguistics*, vol. 14(1), p. 3-33.

MOHAMMED Fuad, 2018, *Social Network Integration and Language Change in Progress in Iraqi Arabic: A Sociophonetic Study of Dialect Levelling in the Hīti Dialect*, Doctoral dissertation, University of Leeds.

OHALA John J., 1981, "The listener as a source of sound change", *in* Carrie S. Masek,. Robert A. Hendrick, & Mary F. Miller (eds.), *Papers from the parasession on language and behavior*, Chicago, Chicago Linguistic Society, p. 178-203.

OUSSANI Gabriel, 1901, "The Arabic dialect of Baghdad", *Journal of the American Oriental Society*, vol. 22, p. 97-114.

RAKIĆ Tamara, STEFFENS Melanie C. & MUMMENDEY Amélie, 2011, "Blinded by the accent! The minor role of looks in ethnic categorization", *Journal of personality and social psychology*, vol. 100(1), p. 16.

ROSE Sharon, 1996, "Variable laryngeals and vowel lowering", *Phonology*, vol. 13(1), p. 73-117.

TAGLIAMONTE Sali A., 2006, *Analysing sociolinguistic variation*, Cambridge, Cambridge University Press.

TALAY Shabo, 2011, "Arabic dialects of Mesopotamia", *in* Geoffrey Khan, Michael P. Streck, Janet C. E. Watson & Stefan Weninger (eds.), *The Semitic languages: An international handbook*, vol. 36, Berlin, Mouton de Gruyter, p. 909-920.

THOMAS Erik, 2010, *Sociophonetics: An introduction*, Basingstoke, Palgrave Macmillan.

THOMAS Erik, KENDALL Tyler, YAEGER-DROR, Malcah, & KRETZSCHMAR William, 2007, "Two things sociolinguists should know: Software packages for vowel normalization and accessing linguistic atlas data", Paper presented at the 36th annual meeting of New Ways of Analyzing Variation (NWAV), University of Pennsylvania, 11-14 October 2007.

VAISSIÈRE Jacqueline, 2009, "Articulatory modeling and the definition of acoustic-perceptual targets for reference vowels", *The Chinese Phonetics Journal*, vol. 2, p. 22-33.

VAISSIÈRE Jacqueline, 2011, "Proposals for a representation of sounds based on their main acousticoperceptual properties", *in* John A. Goldsmith, Elizabeth Hume & Leo Wetzels (eds.), *Tones and features, phonetic and phonological perspectives*, Berlin, Mouton de Gruyter, p. 306-330.

WOLFRAM Walt, 1993, "Identifying and interpreting variables", *in* Dennis R. Preston (ed.), *American dialect Research*, Amsterdam, John Benjamins, p. 193-221.

WOLFRAM Walt, 1993, "Identifying and interpreting variables", *in* Dennis R. Preston (ed.), *American dialect Research*, Amsterdam, John Benjamins, p. 193-221.

Espaces Linguistiques
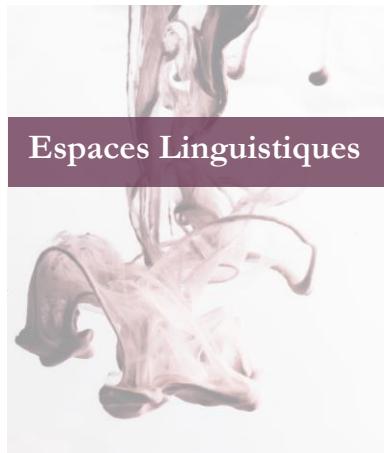
Investigating (in)coherence in Tyneside English(es): sociophonetic variation spectrums of FACE, GOAT, PRICE and MOUTH in the DECTE corpus

Analyse sociophonétique de la cohérence en anglais de Tyneside : les continuums variationnels de FACE, GOAT, PRICE et MOUTH dans le corpus DECTE

Maelle AMAND

University of Limoges, CeReS, UR 14922

maelle.amand@unilim.fr

Abstract: The present study investigates the co-occurrence of FACE, GOAT, PRICE and MOUTH vocalic variants in the Tyneside Linguistic Survey (TLS, Strang, 1968). The analysis is based on the original fine-grained transcriptions of the TLS sociolinguistic interviews across a sample of 44 predominantly working-class speakers (men/women, older, middle-aged and teenagers). Multiple Factorial Analysis and cluster analysis were combined to identify sub-groupings in the variety based on the entirety of the transcriptions (consonants, stressed and reduced vowels), and to investigate coherence across these vocalic variants. The results point towards major gender and class effects with four clusters of speakers characterised by varying degrees of accentedness. Overall, lower-status women and slightly higher-status women distinguish themselves through distinct variant choices. Differences amongst men of either status can best be measured through the extent to which they use the traditional variants.

Keywords: Tyneside English, coherence in sociolinguistics, aggregate approach, Multiple Factorial Analysis, DECTE corpus

Résumé : Cette étude propose une analyse de la co-occurrence des variantes phonétiques des ensembles lexicaux FACE, GOAT, PRICE et MOUTH (Wells, 1982) dans l'Enquête Linguistique de Tyneside (TLS, Strang, 1968). Les transcriptions phonétiques d'origine effectuées sur les dix premières minutes des entretiens sociolinguistiques servent de base à l'analyse, sur un échantillonnage de 44 locuteurs issus de milieux modestes (hommes/femmes, trois classes d'âge). Une analyse factorielle multiple (AFM, Escofier, 1985) suivie d'une classification a permis de classer les locuteurs en sous-groupes en fonction d'une utilisation plus ou moins marquée de l'anglais de Tyneside et d'analyser la cohérence des schémas variationels de ces groupes de locuteurs. L'intégralité du système phonologique disponible à travers les transcriptions a été pris en compte dans l'analyse. Les résultats indiquent que le milieu social ainsi que le genre sont des éléments déterminants dans la classification des locuteurs. Les femmes au statut social légèrement différent se distinguent par un choix plus systématique de variantes vocaliques. La distinction entre les hommes se mesure plutôt à travers la fréquence d'utilisation des variantes traditionnellement associées à l'accent de Tyneside.

Mots clés : Anglais de Tyneside, cohérence en sociolinguistique, approche par agrégat, Analyse Factorielle Multiple, corpus DECTE

Introduction

Tyneside English, henceforth TE, is a variety of English spoken in the north-east of England. The data reported here was collected from the Tyneside Linguistic Survey (Strang, 1968; Strang and Pellowe, 1969). It comprises a series of recorded one-on-one sociolinguistic interviews carried out in the two urban areas of Newcastle and Gateshead (Figure 1), the reading of a wordlist, and acceptability judgements (Allen *et al.*, 2007). Recordings were carried out in the late 1960s and early 1970s. One of the TLS's characteristics is that a portion of the interview material is matched with fine-grained phonetic transcriptions (Beal *et al.*, 2014). A sample of these transcriptions from the recorded sociolinguistic interviews of 44 out of 88 speakers from Gateshead and Newcastle is analysed in this paper–the remaining recordings and their matched phonetic transcriptions were not yet available at the time of the study.

Figure 1: map of Tyne and Wear, north-east England

Source: https://d-maps.com/carte.php?num_car=112548&lang=en

While most studies analysed one linguistic feature at a time (Trudgill, 1972), the TLS pioneered in opting for *aggregate* (Nerbonne & Wieling, 2017) approaches in language variation and change (Pellowe *et al.*, 1972), i.e. the investigation of the relationships among multiple linguistic variables. Nevertheless, the two approaches should not necessarily exclude each other, as Corrigan and colleagues highlight the complementary role of feature-based and aggregate approaches (Corrigan *et al.*, 2014) in revealing various aspects of co-variation in TE. This paper is a compromise between the two approaches: an aggregate statistical analysis is performed on highly enregistered (Agha,

2003) Tyneside vowel variants of FACE, GOAT, PRICE and MOUTH. Jones-Sargent (1983b) performed the first analysis on the fine-grained phonetic transcriptions of the TLS but was faced with the difficulty of a full aggregate analysis due to the computational limitations of statistical programs at the time. Moisl & Maguire (2008) analysed a less fine-grained level of transcription but flaws in the original hierarchical structure of the transcription coding itself led to a major bias amongst certain consonantal features, thereby potentially assigning greater weight to the vowels. The two studies, nonetheless, provided fundamental suggestions for future research: (1) reducing variational noise in the data, i.e. allotting less statistical importance to minor variation patterns, and (2) looking at how stable the clusters across various clustering approaches are. These two approaches were carried out by Moisl (2015) in a later study but with the same semi-fine-grained transcriptions. The present paper draws on the methodological approaches of the two studies: firstly, the use of the fine-grained transcriptions to limit biases caused by the hierarchical structure of consonantal features and secondly, that of noise reduction tools prior to classifying the speakers (cf. methodology section).

However, the innovative methodological aspect of this paper is the application of Multiple Factorial Analysis or MFA (Escofier & Pagès, 1984) to language variation, a tool generally used in sensorimetry (Husson *et al.*, 2001) or in the study of ecosystems (e.g. Lamentowicz *et al.*, 2010). MFA takes into account the hierarchical structure of variational data, namely, the link between a vowel and the variants pertaining to that vowel. It also reveals how linguistic variants covary. Namely, if a speaker uses a regional vowel variant for GOAT, do they use regional variants for the other three vowels? In addition, MFA summarises variation into trends through *dimensions* (ranging from broad to subtle)[44], without obliterating individual variation.

MFA cannot draw boundaries between individuals to form sociolinguistic groups, which is often relevant in language variation. However, when coupled with clustering analyses (Husson *et al.*, 2010), groups with similar variation patterns can emerge. The combined approaches are therefore particularly suited for sociolinguistic research.

Following British sociophonetic conventions for the description of English varieties, the standard keywords FACE, GOAT, PRICE and MOUTH, proposed by Wells (1982, p. xviii-xix), are used to represent the lexical sets containing the four closing diphthongs with the following British Received Pronunciation citation form /eɪ/, /əʊ/, /aɪ/ and /aʊ/. In Tyneside English (TE), words from the

---

[44] Dimensions are also known as factors in a factorial analysis, or principal components in a principal component analysis.

PRICE set often alternate between a raised onset [ɐɪ] and a low onset [aɪ]. This alternation is partially conditioned by what follows the vowel (Milroy, 1996). Milroy's analysis revealed that working-class speakers favoured a raised onset [ɐɪ] in PRICE and that younger male speakers appeared to lead the swing towards a less socially restricted use of this raised variant (Milroy, 1996). Monophthongal realisations for MOUTH [uː] are highly enregistered in Tyneside (Beal, 2004)–i.e. the general public is aware that this vowel is typical of Tyneside. But for most speakers, it is often used for a limited number of words like *town* or *brown*, respectively spelt *toon* (the *Toon Army*)[45] and *broon* (*Westmorland Broon ale*), to reflect the local pronunciation. Beal (2004) states that the most widespread realisations of MOUTH in Tyneside are closer to [ɛʊ].

FACE and GOAT are, to some degree, partner vowels working in *lockstep* since "the ways in which they vary in terms of their phonetic exponency and the distribution of these variants across the TE-speaking population seem to work in parallel" (Watt & Foulkes, 2017, p. 157). The most commonly used variants are [eː] and [oː] for FACE and GOAT, respectively (Watt, 1998). They are said to be unmarked in the area because they "accommodate a less marked identity as Northerners" (Haddican *et al.*, 2013). They will be referred to as the *pan-northern variants* (*Ibid.*). The local variants are often realised as forms approximating [ɪə] for FACE and [ʊə] or [øː] for GOAT (Corrigan, 2012).

Closing diphthongs [eɪ] and [oʊ] are often heard among middle-class speakers (Watt & Foulkes, 2017) and will be call *prestige-variants*. Given the asymmetry across the variant inventories for each vowel and the social indexicality observed in more recent TE speech (Watt, 1999; Milroy, 1996), the working hypothesis of this paper is that coherence in TE is only partially the result of linguistic constraints such as the Scottish Vowel Length Rule (Milroy, 1996) or the type of phonemes on either side of FACE and GOAT vowels (Buchstaller *et al.*, 2017) and that social factors, first and foremost, play a major role in accounting for correlation in usage–as found in many sociolinguistic studies on other varieties (starting from Labov, 1972; Macaulay 1977 and Trudgill 1974). However, studies investigating co-variance (*coherence*) in a variety are rarer (cf. Guy, 2013; Beaman & Guy, 2022). This was also amongst the original aims of the Tyneside Linguistic Survey (Jones-Sargent, 1983, p. 21-22). Following De Camp's study in 1971 on Jamaican English, more recent approaches have delved into the blurry boundaries of sociolinguistic stratification, thereby looking at

---

[45] The local football club.

continuums instead of discrete groups (e.g. Amand, 2019 on TE, Sewell, 2022 on Hong Kong English consonants).

The purpose of this paper is to report intercorrelations of four highly enregistered (Agha, 2003) TE vowel variants, i.e. those of FACE, GOAT, PRICE and MOUTH based on the original TLS phonetic transcriptions. The following questions are raised: (1) **Lects**: how many linguistic groups can be found through the variant choice of FACE, GOAT, PRICE and MOUTH only? (2) **Coherence**: do these groups use variants consistently across FACE, GOAT, PRICE and MOUTH in TE? (3) To what extent are the lects, and the coherence within, socially motivated? (4) Can variation patterns be defined less in terms of discrete lects, but rather in terms of multiple variational **spectrums**?

1. Methodology: the corpus

1.1. The TLS phonetic transcription scheme

The data analysed in this paper is referred to as the TLS-coding data. It is composed of fine-grained phonetic transcriptions (carried out by trained phoneticians) of the first ten minutes of recorded sociolinguistic interviews. The recordings were made in the participants' homes. These transcriptions comprise consonants, stressed and reduced vowels. To make the transcriptions computer readable, each phonetic variant deemed likely to appear in speech was encoded as a sequence of five digits (Jones-Sargent, 1983, p. 295-302), incorporating a two-level hierarchical structure (Figure 2). The full coding-scheme is available on the DECTE website (Jones-Sargent, 1983a)[46].

Figure 2: original transcription cards (left) & TLS coding-scheme (right)

---

[46] https://research.ncl.ac.uk/decte/transscheme.htm
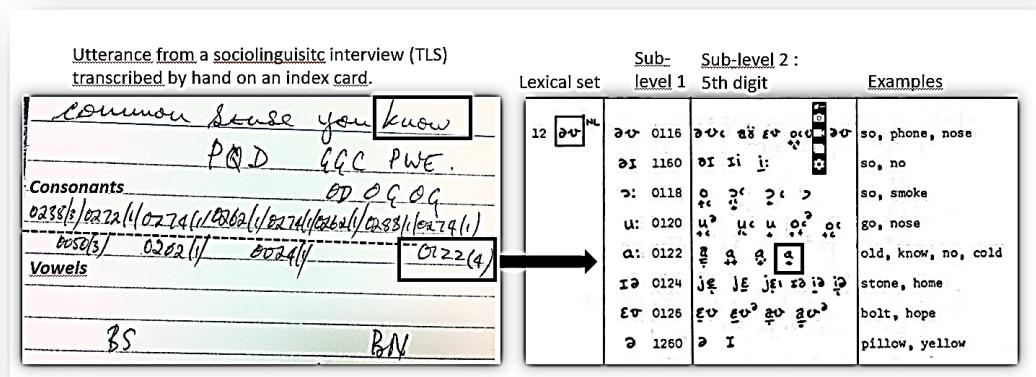
1.2. Phonetic data

The sample of the TLS-coding data analysed is composed of frequencies of 566 variant types distributed across all 63 phonetic features (consonants and vowels) during the first 10 minutes of a sociolinguistic interview for 44 speakers. The number of variant types across the 4 vowels amounts to 84. The total number of transcribed variants amounts to 105,204 tokens, 9,428 of which pertain to the FACE, GOAT, PRICE and MOUTH lexical sets. The median number of occurrences per speaker for the latter four features ranges between 21 and 57. Frequencies were computed from raw counts of all variants within each lexical set.

Table 1: counts per speaker per vowel (total count: top row, speakers: n= 44).

SUM FACE	1 705	SUM GOAT	2 515	SUM PRICE	4 186	SUM MOUTH	1 022
SPEAKER MEAN	39	SPEAKER MEAN	57	SPEAKER MEAN	95	SPEAKER MEAN	23
SPEAKER MEDIAN	36	SPEAKER MEDIAN	57	SPEAKER MEDIAN	99	SPEAKER MEDIAN	21
MIN	16	MIN	11	MIN	19	MIN	6
MAX	75	MAX	107	MAX	151	MAX	55
Q1	29	Q1	46	Q1	82	Q1	17
Q3	48	Q3	70	Q3	110.75	Q3	29

1.3. Selected social data

The original aim of the TLS was to include as many social variables as possible–father's occupation, parents' birthplace, political affiliation, leisure, identification with the area, and so on, in order to build a hierarchy of the social characteristics that were more significantly linked to the linguistic variational patterns (Jones-Sargent, 1983). In the present study, we chose to retain macrosociological categories (Eckert, 2012) such as age, gender and social class. A fourth category regarding whether speakers had continued their education after the minimum school leaving age was also included.

1.4. Participants

We selected the 7 Newcastle and 37 Gateshead participants from the original TLS study (see Table 1) whose data had been checked by the DECTE team, and made available on the DECTE website as of 2014. Since these speakers were randomly selected from the Electoral Register of 1966 by the "rateable value per dwelling by polling district" (Pellowe *et al.*, 1972, p. 23-24), the number of participants within each macrosociological category, i.e. age, gender and class, is not necessarily even across all categories. The class categories are based upon the speaker's (former) occupation. The *Classification of Occupations* of 1966 served as a reference to better fit with the social categories defined by the original TLS team in the 1960s-70s.

Table 2: Social background of the 44 TLS informants

Lower-middle class[1]	Women (n=6)	Men (n=5)
	Civil servant, school secretary, nurse, employment exchange worker...	Skilled workers (engraver, cast operator), teacher trainee...
Young (F=3, M=3)	G05F, G09F, G21F	G20M, G23M, N04M, N02M, N06M
Middle-aged (F=2, M=2)	G11F, G34F, N01F, N03F	--
Older (F=1, M=1)	--	G28M

Working class[2]	Women (n=15)	Men (n=13)
	Home help, factory worker, cleaner, waitress, shop assistant...	Painter, plumber, storeman, miner...
Young (F = 3, M = 2)	G08F, G10F, G15F	G12M, G27M
Middle-aged (F=12, M=10)	G01F, G03F, G16F, G17F, G22F, G25F, G26F, G32F, G35F, G36F, G37F, N05F	G02M, G04M, G07M, G14M, G18M, G19M, G29M, G31M, N06M, N07M(?)[3]
Older (F=1, M=4)	G06F	G04M, G13M, G24M, G33M

Middle-aged speakers predominate. We therefore expect a certain degree of *retrenchment* (Chambers, 2003, p. 195; Wagner, 2012, p. 375) from the use of more local or less-standard features used in youth. Two speakers were not born in the North-East (G31M: Northern Ireland, N06M: south-east London) but were included in previous studies (Moisl, 2015; Jones-Sargent, 1983b). In our study, these *outliers* serve as points of reference in the statistical analysis to highlight the distance between native Tynesiders and speakers who have settled in the area but were not born and raised in the North-East. They contribute to building a limit in the variational space computed by the statistical analysis, whose first aim is to simplify complex data and show macro-groupings of individuals (native Tynesiders vs. non-native Tynesiders) before going into meso-groupings within the native Tynesiders, e.g. broad vs. less broad.

The next section details the statistical methods applied to the corpus.

2. Methodology: statistical analysis

2.1. Multiple Factorial Analysis + clustering in language variation

The statistical tool used here to analyse the data is rarely used in sociophonetics. The methodology of the approach along with the interpretation of the results will therefore be detailed in greater length than is commonly found in articles on language variation.

A Multiple Factorial Analysis was carried out with the R package {FactoMineR} (Lê *et al.*, 2008). It was selected over other statistical methods because it is *feature-centric* (Amand, 2019). Namely, each variant (or *exponent*) is not treated statistically as an independent variable but as a variable that is hierarchically linked to its pertaining phoneme or *feature*[47]. Hence, the results are first apprehended from the perspective of the *feature* itself, e.g., how variation in FACE helps define speakers socially.

Unlike more common methods such as Principal Component Analysis (PCA), MFA considers the structural relationship between the linguistic features and their variants. Since there are over 20 different variants reported for FACE and GOAT (see Amand, 2019), a *variant-centric* approach with PCA would have generated output that is difficult to interpret because each variant is treated as an independent variable and not as pertaining to a lexical set. The approach is similar to the one used in Ghyselen and De Vogelare (2018) who computed speaker-profiles based on correspondence analysis, which also allows for sublevels of a main variable.

Another reason for using MFA is that it gives less frequent vowels like SQUARE (RP /eə/) or vowels with more complex variation "a chance to have the same importance in the final analysis" (Picardi *et al.*, 2020) as it balances out the role of each lexical set in accounting for variation. In a canonical PCA analysis, features with binary variation such as *-ing* (velar /ŋ/ or apical /n/) will appear as overly important in the sociolinguistic discrimination of speakers. Namely, the easier it is to split speakers (those who favour /ŋ/ vs. /n/), the more salient the lexical set will appear in the analysis. The other features, with more complex variation, will be minimised (Escofier & Pagès, 1994). Since the number of reported variants are not the same across the four vowels under scrutiny, MFA was deemed a more suitable approach than PCA.

---

47 The term feature is used here because the approach is not restricted to phonetic variation and can be applied to other types of variation (lexical, grammatical etc.).

A cluster analysis was then performed based on the MFA output. This provides an ideal combination: (1) MFA prunes variational noise (i.e. a variant which occurs only once across two speakers) without downplaying the role of relevant variation patterns, and (2) the cluster analysis builds on the MFA analysis to identify more stable sub-groups of speakers (Husson *et al.*, 2010). The Ward method, often also called the 'minimum variance method,' was used for clustering since it has proven relevant for the TLS data analysis (Moisl, 2015) and in several other linguistic studies (Deumert, 1999; Gries, 2009). It also minimises intra-cluster variance (Janssens *et al.*, 2008; Gries, 2009) so as to limit the number of groups containing one or two speakers (see Amand, 2019 for a comparison of clustering methods on the full TLS data). Finally, I evaluated whether these groups shared social characteristics using chi-square tests provided by {FactoMineR}.

2.2. Preliminary analyses: assessing the robustness of the MFA results

Preliminary analyses confirmed the lack of randomness in the data (Husson *et al.* 2011, p. 229) in variation and the need to favour MFA over other methods like PCA (see Amand, 2019, p. 345 ff.). The MFA analysis was first performed on the dataset with all 566 phonetic variants from the 63 phonological features (consonants, full vowels and reduced vowels), so as to check whether the variants of FACE, GOAT, PRICE and MOUTH have comparable discriminating roles within the context of the original 63 phonological features retained in the TLS coding-scheme (see Amand, 2019, p. 355).

In MFA, factors, aka. *dimensions*[48], have to be viewed in terms of axes with opposite variational patterns at each end. Another element to consider is the presence of diverging variant usage, with specific variants emerging as representative of opposite variant usage. These variants play a major role because they define the limits of the variation spectrum which corresponds to a dimension. The first five variables that contribute the most to the definition of the first two dimensions percentagewise are examined. The higher the score, the better the variants help define opposite trends between speakers. Given the high number of variables included in the model, these percentages are necessarily low. But if the 84 variants had contributed to the dimensions evenly, the percentage would have been 1.19%, i.e. 84/100, while the top contributions are between three to eight times as high. Contributions were deemed reliable enough.

Another rule of thumb is to check whether eigenvalues are superior to 1. Here, eigenvalues drop below 1 from the sixth dimension onwards, (cf. Baayen, 2008) and the first six dimensions have

---

[48] In the "French School of PCA" (Holmes, 2006), factors are more generally called dimensions.

eigenvalues ideally superior to 1 (Di Franco & Marradi, 2013), which means that the first five dimensions are relevant in measuring variation across the sample of speakers. The remaining ones are potentially noise or marginal variation patterns. However, only the two dimensions, with respective eigenvalues of 2.81 and 1.86, will be used to plot the clusters on the factor map, since visualising six dimensions simultaneously leads to a reduced comprehensibility of the results (Ghyselen & De Vogelaer, 2018). After an inspection of the clustering output results, a maximum of nine groups was retained for the classification process because a lower number of groupings would have ironed out sub-groups using rarer variants.

3. Results

3.1. An overview of the main variation patterns

Aggregate percentages (Figure 3) provide a general overview of dominant forms in Tyneside in the 1970s in the given sample of 44 speakers. In FACE and GOAT, the pan-northern monophthongal realisations in yellow predominate (FACE: 59%+2%=61%, GOAT: 44%+7%+3%=54%). Prestige forms with closing diphthongs (in blue) have similar scores (respectively 10%+3%+2%=15% vs. 16%+2%=18%), so do localised variants (17% vs. 21%). GOAT and MOUTH have a wider range of localised variants (green) with various monophthongal and diphthongal variants. The highly enregistered monophthong for MOUTH (as in *Toon*) is indeed extremely rare in casual interview speech (7%).

Figure 3: percentage frequency per variant per vowel for all 44 speakers
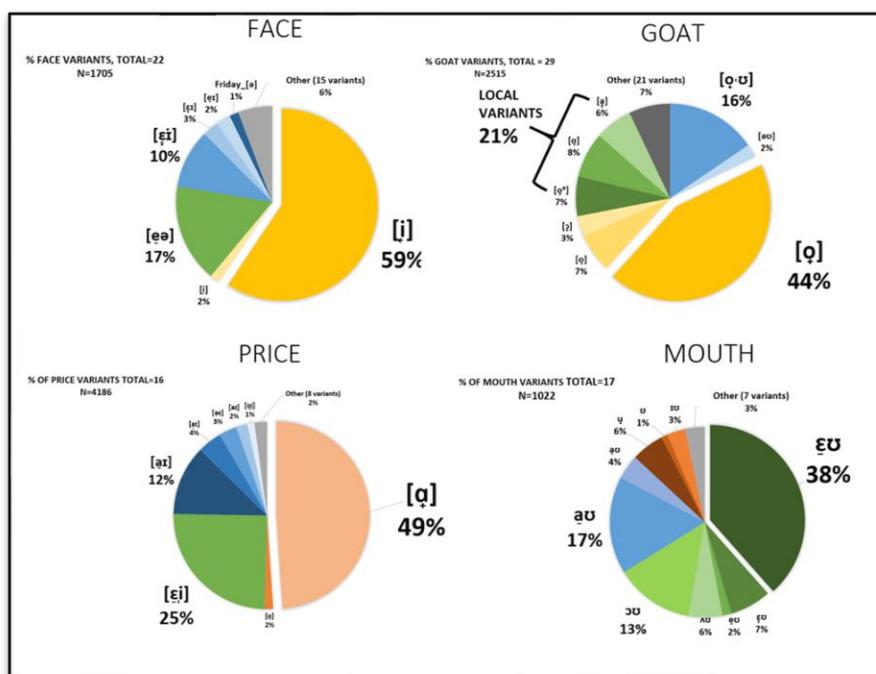
28 speakers use the pan-northern variant [eː] for FACE over half the time, but only 20 use [oː] for GOAT above 54% of the time, this means other variants compete with the pan-northern one. 17 speakers use the raised onset [ɛʊ] in MOUTH above 50% of the time. Scores for the raised onset [ɛɪ] in PRICE are lower due to the inclusion of the personal pronoun *I*, often realised as [ɑː] (Moisl & Maguire, 2008) but 20 speakers use the raised onset around three times as much as the low onset.

3.2. "Are yi broad?": speakers at opposite ends of 2 variation spectrums

Interpreting the dimensions: here, dimensions reflect spectrums where speakers are placed depending on their overall use of variants across the 4 vowels. Interpretation can be made easier through a plot with two axes crossing at a 90-degree angle (Figure 4). Speakers with completely different variation patterns are placed at opposite ends of that spectrum. Dimension 1 generally displays speakers on an obvious opposition: those who use non-local vs. those who use local forms. Dimension 2 shows a slightly subtler pattern opposing speakers at either end. Namely, amongst those who do not use standard forms, sub-variational trends emerge: users of vestige forms vs. users of pan-northern forms. Other dimensions show micro oppositions amongst the pan-northern speakers or amongst the traditional speakers.

Figure 4: Schematization of the first two dimensions of MFA, used to highlight major and subtler variation patterns in sociophonetics
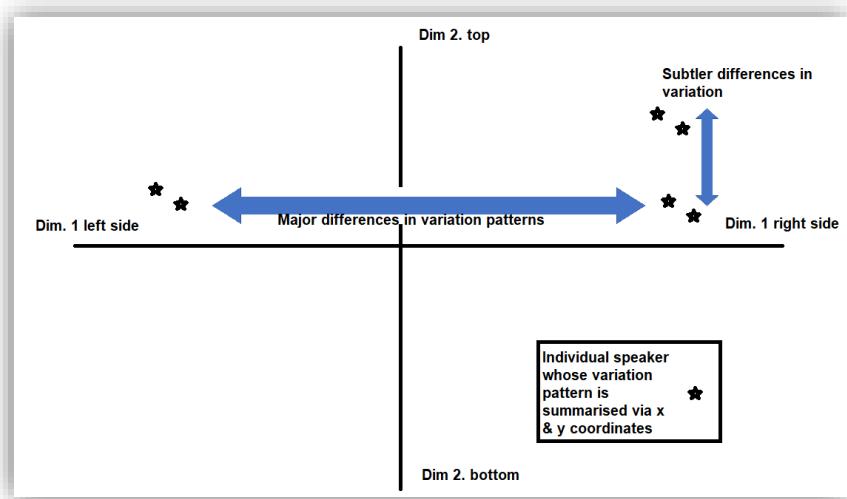
Oppositions in variant use through dim. 1 & 2: the MFA diagnostics for dimension 1 (Table 2) highlight an opposition between the speakers who exhibit a very homogeneous variation pattern in all four vowels (the more socially prestigious forms), and the speakers who use the regional variant [ɛɪ] in PRICE, but who do not necessarily use the same variants in FACE, GOAT and MOUTH.

This means that there are two main groups: one who uses standard forms consistently across the 4 vowels, and another group whose speakers commonly favour [ɛɪ] in PRICE significantly, but who distinguish themselves through the use of distinct regional forms. The opposition within the latter group is reflected in dimension 2, which reveals an opposition between the speakers who use the supralocal variants [eː] and [oː] in FACE and GOAT, plus the raised onset in MOUTH ([ɛʊ]), and those who exhibit the traditional variants of PRICE [i] and MOUTH [u].

Table 2: Top 5 contributions of the phonetic variants to dimensions 1 (top) and 2 (bottom) along with the position of the variants on the dimension axes when variant use is above average

Dim. 1 — Above supralocal vs. regional speakers

Variants at opposite ends of dimension 1 (top contributions in %)			Side of the axis on factor map when score is high
PRICE	[ɛɪ]	7,83	Right (positive coordinates)
FACE	[eɪ]	5,19	Left (negative coordinates)
GOAT	[oʊ]	4,97	Left
PRICE	[aɪ]	4,6	Left
MOUTH	[aʊ]	4,08	Left

Dim. 2 — Supralocal vs. traditional speakers

Variants at opposite ends of dimension 2 (top contributions in %)			Side of the axis on factor map when score is high
GOAT	[oː]	9,77	Bottom (negative coordinates)
FACE	[eː]	6,22	Bottom
MOUTH	[ɛʊ]	5,97	Bottom
MOUTH	[u]	3,71	Top (positive coordinates)
PRICE	[i]	3,3	Top

Summary of both tables

Dim. 1: (a) if you had to find one vowel that distinguishes typical Tyneside speakers from pan-northern speakers, take PRICE and look for the raised onset; (b) when a speaker uses pan-northern forms, they seem to use them across the four vowels; (c) within the typical Tyneside speakers, there is a high degree of variation and sub-groups can be found through dimension 2.

Dim. 2: (a) across typical Tyneside speakers two trends stand out: speakers with high scores of pan-northern forms in GOAT, FACE and MOUTH vs. speakers with vestigial forms in PRICE and MOUTH.

However, these covariational trends need to be confirmed through an inspection of the correlation between the variants, and groups of speakers, confirmed via a cluster analysis. This is dealt with in the next section.

3.3. Coherence across the 4 vowels and across speaker groups

Covariation patterns will now be examined in more depth via a correlation circle displayed in Figure 5 (left). Correlation circles are particularly useful for visualizing covariation amongst a high number of phonetic variants in two dimensions simultaneously. Variants are depicted in the form of arrows. The closer the arrows are to the rim of the circle, the better they reflect a clear opposition between speakers in one or both dimensions (Abdi and Williams, 2010). To ensure a better interpretability of the graph only the 12 phonetic variants with the highest contributions to dimensions 1 and 2 are labelled here[49]. Arrows forming acute angles are positively correlated (FACE [eə] & MOUTH [u:]). Obtuse angles indicate a negative correlation (raised vs. low onsets in PRICE). Variants with no correlation form a right angle (GOAT [a] vs. FACE [e:]).

Three main covariation patterns emerge: the top left part of the circle displays covariation amongst high prestige variants and the top right, amongst traditional ones. The bottom part of the graph shows a more supralocal type of variation with a raised onset in MOUTH and the pan-northern [e:] and [o:] in FACE and GOAT. The arrow for the raised onset in PRICE [ɛɪ] is placed in between the pan-northern and traditional arrows. This means that many speakers from both groups use this raised onset.

Figure 5: correlation circle (left): coherence in variants & factor map (right): individual speakers their affiliation to a cluster
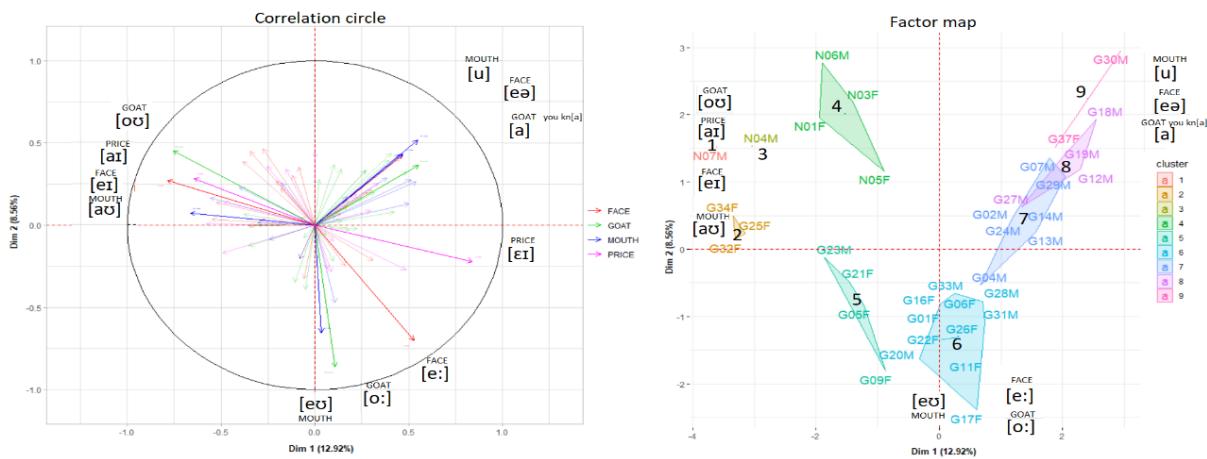
The correlation circle confirms the output for dim. 1. Prestige variants covary: speakers who use closing diphthongs in GOAT and FACE, use low onsets for both PRICE and MOUTH. The traditional

---

[49] There were two monophthongal variants for MOUTH but [u] was used to label them both (TLS code: 01462 and 01464).

variants that covary in a highly systematic way are the centring diphthong [eə] in FACE and two monophthongal realisations in MOUTH ([u]). The variant [a] for GOAT is also positively correlated with the traditional variants [eə] and [u]. The raised onset in MOUTH [ɛʊ] is strongly correlated with the pan-northern variant [oː] for GOAT. The pan-northern FACE variant [eː] is less strongly correlated with [ɛʊ] and [oː] because many traditional speakers use it too. This is also the case for the raised onset in PRICE. The correlation circle, however, does not show individual speakers and limits itself to coherence amongst variants. Another graph with the same coordinates needs to be added to reveal the distances between each speaker.

The right part of Figure 5 illustrates the distances between individuals and is called a *factor map*. This graph is the counterpart of the correlation circle and the two should ideally be interpreted side by side since they represent the same spectrums represented by dim. 1 and 2, albeit highlighting different aspects of the data, which is similar to the use of different filters on the same map. While the correlation circle highlights coherence within the variants chosen by speakers, the factor map places the speakers based on their overall variant usage. On the factor map, speakers with high percentage scores in one of the above-mentioned variants will be placed around where variants' arrows are pointing in the correlation circle. Speakers with similar phonetic variation have similar coordinates (G22F and G26F). By contrast, G32F, G17F and G30M have very distinct pronunciation patterns. The closer the speakers are to where axes cross zero, the closer they are to the sample's average variation pattern, e.g. G33M and G04M.

Polygons circumscribe speakers that were grouped together in the cluster analysis. The coordinates of certain speakers suggest considering at least more than 6 groupings and setting apart atypical speakers that are placed far away from the rest of the sample (N07M or N04M). We now comment on the results of the cluster analysis.

3.4. Sub-linguistic norms in Tyneside English?

How many dimensions retained for the clustering analysis? In order to analyse the clustering tendencies in more depth, the first five dimensions of the MFA were used as input for a cluster analysis–the eigenvalue for the sixth dimension was 1.01 and was disregarded as it was too close to the rule of thumb threshold of 1. The analysis was consolidated with k-means, so as to optimise group homogeneity (Lebart *et al.*, 2000). After inspecting the sum of squares, it was decided to stop at the last large drop in units, namely, between 8 and 9 clusters, in order to have a better idea of

the variational cohesion within and between these sub-groups. Going beyond that point would have led to overfitting (cf. Baayen, 2008).

Figure 6: classification of speakers. First letter of speaker index: N = Newcastle, G = Gateshead. Last letter: F = woman, M = man. Groupings are the same as Figure 3. Y-axis: the higher the node, the more different the overall variant patterns. X-axis: cumulative sum of speakers (n=44)
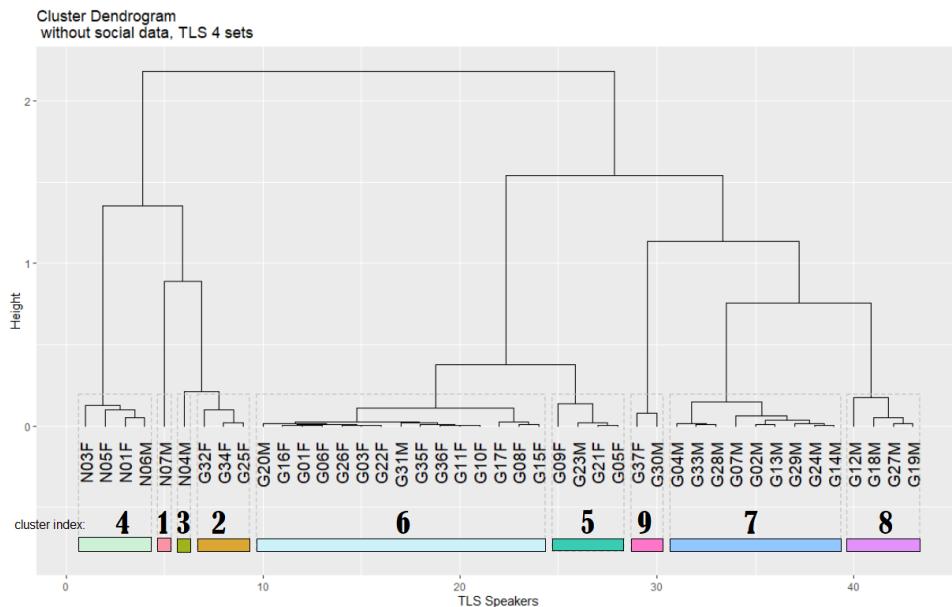
What are the linguistic groups of speakers? Figure 6 illustrates the dendrogram resulting from the classification. It indicates that the speakers are divided into two main categories (cut-off point superior to 2, y-axis): the Newcastle speakers and all but three of the Gateshead speakers (G32F, G34F and G25F). The higher the split, the larger the difference between the speakers and the lower the split, the closer the speaker's values approach the cluster's average scores. Based on the MFA plots (Figure 3), we know that this reflects the opposition between the speakers with a higher share of prestigious variants and those exhibiting higher rates for the regional variants (dim. 1). The second major split (cut-off point of 1.5) is the one between the more traditional speakers and those who predominantly use pan-northern variants (dim. 2). Beyond oppositions defined in dim. 1 and 2, the speakers who use the prestigious variants can also be sub-divided into two categories, with N07M being considered an outlier. **There are also three categories of traditional speakers and two for pan-northern ones**. We now examine which clusters are correlated with one or more social factors.

How do these groups correlate with social factors? Chi-square tests provided in the MFA + clustering output tables assess the link between the linguistic groupings and the social factors, here: class, age, gender, education and sampling area (Newcastle vs. Gateshead cohorts).

The selected percentages given in this section are also found in the raw output tables. With the default level of group cutting, (i.e. 6 sub-groupings) one major effect is the city of residence (p = .009). This is partly due to variations in phonetic transcriptions[50], and, potentially, to class[51]. Yet, the social factor that best helps discriminate speakers is gender (p = .003). This effect is clearly visible in Figure 4, with most speakers being grouped by gender. This suggests that in Tyneside English, there are distinct male and female linguistic norms reflected in the phonetic variants of FACE, GOAT, PRICE, and MOUTH. With the 9 sub-groupings[52]: class, sex and city of residence significantly contribute to define at least one of these 9 groups (respectively p. < 000003, p < .003 and p < .03). Let us now examine what characterises these groups from a socio-linguistic point of view.

Cluster 6 includes more speakers (n=14). It is predominantly female (p < 0.05, 13 women vs. 2 men) who did not pursue further education beyond minimum school leaving age (p < 0.05, 88.2% with no further education). This group extensively opt for the pan-northern variants [oː] and [eː] for GOAT and FACE and favour the raised onset in MOUTH [ɛʊ] (above 75% for the three variants). Both local *and* high prestige variants are avoided in this group as they remain below 6%.

Pan-northern female speakers → monophthongs for GOAT and FACE, raised onsets for PRICE and MOUTH. → highly coherent across the 4 vowels.

By comparison, cluster 5 is exclusively composed of younger speakers from lower-middle class background and with further education (class: p = 0.003; education: p = 0.004; age: p = 0.004)[53]. Clusters 5 and 6 mainly differ in variant usage for MOUTH and PRICE. While women with no further education in cluster 6 favour the raised onsets for both these vowels, the group of younger lower-middle class speakers used low onsets above 90% for MOUTH (except for G09F who uses the raised onset more than 60%). However, both groups predominantly use the pan-northern variants [oː] and [eː] for GOAT and FACE (except for G21F who has an equal share of prestige and pan-northern variants in FACE).

---

50 There were two distinct phonetically-trained transcribers from Newcastle University. Although both followed the same transcription schemes, minor differences in fine-grained transcriptions were observed. A more careful examination of these differences will be carried out in a future study.

51 This issue is being investigated further since more information is being made available on the Newcastle speakers.

52 The clusters are consolidated with k-means. This parameter has an impact on how clusters are made and on the final diagnostics.

53 Although the group includes a majority of women (3 women vs. 1 man), the group is too small for gender to be deemed a significant effect.

Lower-middle class younger speakers → monophthongs for GOAT and FACE, raised onsets for price but low onset for MOUTH. → non-localised speech in MOUTH only.

Clusters 7, 8 and 9 are almost exclusively composed of men, except for G37F in k9. Clusters 7 and 8 pattern in a similar way for FACE (37.1% vs. 38.3%) and PRICE (37.6% VS. 44.0%) but cluster 8 is characterized by the significant, albeit infrequent use of the vestigial variant [u] in MOUTH (18.3%).

In addition, two separate retracted MOUTH variants seem to define these two groups–cluster 7 [ɔʊ] 45.0%, cluster 8 [ʌʊ] 34.5%. An acoustic analysis along with a perception test should be carried out to assess the differences between the two variants. Cluster 9 groups two atypical speakers (G30M, G37F) who, combined, use [u] for MOUTH twice as much as cluster 8 (53.0% vs. 18.3%). These speakers also used significant amounts (ca. 11%) of rare variants like [ɪi] and [ɛʊ] for GOAT, and [ɪʊ] for MOUTH.

Speakers using highly localised forms are predominantly men (except for G37F). But some speakers distinguish themselves through the use of [u] in MOUTH (K8 & K9), while others use retracted onsets [ɔʊ] or [ʌʊ] (k7 & k8). Working-class men form a continuum from broad to very broad (k7 → k8 → k9).

Cluster 4 is exclusively composed of speakers from the Newcastle sample. It is defined by particularly high scores of [oʊ] in GOAT (65.6%) and two competing variants for FACE (pan-northern 11.5% vs. prestige form 17.8%). It has a large share of raised onsets in MOUTH (43.3%) and 10% of low onsets for PRICE. **Coherence in this group is less obvious because it is in between two linguistic profiles: pan-northern speakers and those using prestige forms.**

By opposition, cluster 2 uses the low onsets in MOUTH and PRICE above 80%. This group also used the prestige variant [eɪ] in FACE more extensively than cluster 4 (k2: 74% vs. k4: 17.8%). It does not quite work in lockstep with GOAT ([oʊ] in 25%). Cluster 2 is only composed of three Gateshead women and none of the selected social variables significantly characterize the cluster due to its size. We hypothesise that **cluster 2 forms the middle part of the female variation continuum ranging from non-localised (k4,1,3) to female pan-northern (k6 & k5).**

4. Discussion

In the Gateshead sample, variation is highly determined by gender: men use variants co-occurring amongst men only. The same is found for women, with a few rare exceptions (G37F, G20M). The

broad⇔less broad spectrum for Gateshead men ranges from pan-northern forms to vestigial forms. Women's spectrum (except for G37F) goes along the line of pan-northern⇔prestige forms. As expected, the common ground is pan-northern features used by both men and women, which is supported by the work of Haddican et *al.* (2013).

In prestige and pan-northern variants for FACE and GOAT, co-variation is coherent whereas, for local variants, co-variation was harder to assess due to the wide array of variants reported for GOAT compared to those for FACE, which was mainly [eə]. In addition, the more frequent use of pan-northern [e:] for FACE amongst the traditional speakers indicate a looser symmetry between FACE and GOAT in the masculine and traditional sub-varieties of TE.

PRICE and MOUTH also co-varied tightly amongst most speakers using supralocal and high-prestige variants for FACE and GOAT. While users of supralocal forms for FACE and GOAT favour a raised onset for both PRICE and MOUTH, speakers using high-prestige variants for FACE and GOAT can be divided into two sub-groups, i.e. those who use low onsets for both PRICE and MOUTH and those using a low onset for MOUTH but a raised onset for PRICE. The latter group corresponds to younger lower-middle class women who benefited from further education. Amand (2019, p. 513 ff.) investigated the interactional effect of class and age for the raised onset in MOUTH, which confirmed that, in the 1970s, the variant simultaneously indexes class and age in TE–young working-class women use the variant twice as much as their lower-middle class female counterparts (respectively > 80% vs. > 40%). However, this variant did not seem to index class amongst older women whose scores range between 70% and 80%. It is beyond the scope of this paper to elaborate upon age effects but it is possible that this age effect implies standard change with the low onset in MOUTH becoming the new local standard amongst women of both classes as the latter variant is more often used in read speech in the TLS and even more so in the read material of a 1990s corpus of TE (Amand, 2019). It is also possible that the low onset in MOUTH is simply an extension of the pan-northern coherence observed in GOAT and FACE to the mouth vowel.

Raised onsets in PRICE co-vary in a peculiar way since its covariation with pan-northern variants in GOAT and FACE and the raised onset in MOUTH does not prevent a simultaneous co-variation with the traditional variants of FACE, GOAT and MOUTH. This indicates that in the 1970s, a raised onset in PRICE indexed class more than it did gender. These results are in line with Milroy's study on a later corpus of TE (Milroy, 1996). It may also index an intermediate degree of localness encompassing both users of traditional forms of TE in other vowels and speakers with a more levelled form of TE in their use of pan-northern variants in FACE and GOAT.

One regular pattern observed here is that the working-class women of the sample systematically score higher in the pan-northern variants of FACE and GOAT, which is in line with Watt's study of conversational material from the TLS corpus (Watt, 1998). They nonetheless retain raised onsets in PRICE and MOUTH; the former indexing class or localness, while the other seems to index both class/localness *and* gender. In addition, women form more coherent sociolects than men because their choice of variants for each vowel is more limited, which converges with Guy's findings that women, contrarily to men, "consistently show co-varying usage of socially stratified variables" (Guy, 2013, p. 69). Nonetheless, men are also coherent in the sense that they systematically avoid the variants adopted by the less traditional women.

Interestingly, G20M, is an atypical speaker concerning MOUTH. He was the only man using the record score of 90% for the raised onset in MOUTH vs. an average of 36% amongst the other men. Buchstaller *et al.* (2017) analysed his realisations for FACE 42 years later, in a second sociolinguistic interview, in similar conditions to the original TLS interview. At the time of the first interview, he was a teacher trainee in his twenties and his FACE variants were almost categorically monophthongal. Given the increasingly high number of women trained as teachers in the 1970s (Worth, 2019)[54], It is possible that G20M used the raised onset in MOUTH as a form of accommodation to his colleagues at work (cf. Coupland 1980 on style shifting in the work-place), a habit that was triggered by the semi-formal setting of the sociolinguistic interview, also known as the *observer effect* (Labov, 1972 and 1981). In the second interview, he had experienced considerable upward mobility and his realisations of FACE shifted towards "the most standard repertoire in the panel sample" (Buchstaller *et al.* 2017, p. 24), namely, the closing diphthong [eɪ]. G20M works in a language sensitive occupation (Chambers, 1994), in which the "legitimised language" plays an important role in shaping variation patterns amongst speakers, which may account for this remarkable shift. It is hypothesised that this change in FACE in favour of a more southern-like form is likely to be observable in MOUTH and PRICE too, with a complete retrenchment from the use of the raised onsets for both vowels, and higher ratios of closing diphthongs for GOAT. A future study on coherence in FACE, GOAT, PRICE and MOUTH on the second interviews of the six original 6 TLS speakers may provide insight as to how coherence evolves across the lifespan.

---

[54] The 1970s were a turning point for female educators. As Worth explains, "there was a marriage bar in place across much of the public sector until the 1940s and, crucially, part-time or more flexible roles were not widely available in the welfare professions until the 1970s". She adds that "in the 1971 census, 381,000 women were studying under subjects classified as 'education', compared to only 140,000 men".

Conclusion

This paper explored sociolectal cohesion within Tyneside English. In the Gateshead sample, four groups emerged as being sharply stratified by gender and social status. Although, at first sight, women seem more coherent in their variant use, men are also coherent in their own way. Women may treat socially stratified variants as discrete entities, i.e. opting for either one variant or another distinct one, whereas men seem to treat socially stratified variants as a continuum, thereby distinguishing themselves through the degree in which they exhibit traditional variants.

Although certain linguistic groups seemed less coherent than others with regards to the four vowels under scrutiny, the average standard deviation for groups of more than two speakers ranged between 6 and 11%. The holistic model, "in which particular varieties or registers characterised by clusters of socially evaluated variables become identifiable in a community" (Guy, 2013), is supported by the fact that a certain number of speakers have MFA coordinates that approximate the centroids of a cluster in a factor analysis. They are what Pagès (2013) calls *paragons*, or the speakers with the average characteristics of the cluster they belong to, e.g. the cluster of "Geordie[55]" men: G14M, G13M, and G24M (respectively, former factory worker, former miner and former wire drawer, aged between 40 and 60). This means that these paragons' variation patterns can form the basis of holistic representations of sociolinguistic sub-groups, with, at the periphery of clusters, innovators of sound change, social risers (G20M, the teacher trainee) and speakers that keep vestigial forms alive (G19M, the JCB driver or G37F, the factory worker).

References

ABDI Hervé & WILLIAMS, Lynne. J. 2010. "Principal component analysis", *Wiley interdisciplinary reviews, Computational statistics*, vol 2, p. 433-459.

AGHA Asif, 2003, "The social life of cultural value", *Language and Communication*, vol 23, p. 231-73.

ALLEN Will, BEAL Joan C., CORRIGAN Karen. P., MAGUIRE Warren. and MOISL Herman L., 2007, "A Linguistic 'Time Capsule': The Newcastle Electronic Corpus of Tyneside English", in BEAL Joan C., CORRIGAN Karen P. & MOISL Herman L. (dir.), *Creating and digitizing language corpora: Volume 2: Diachronic Databases*, London, Palgrave Macmillan UK.

---

[55] The term Geordie refers to people from Newcastle and is often more closely associated with working-class men (Beal, 2009).

AMAND Maelle, 2019, *A sociophonetic analysis of Tyneside English in the DECTE corpus: The case of FACE, GOAT, PRICE and MOUTH* [online], Sciences du Langage, Université de Paris Diderot / Newcastle University, available at: https://u-paris.fr/theses/detail-dune-these/?id_these=4248

BAAYEN R. Harald, 2008, *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*, Cambridge, Cambridge University Press.

BEAL Joan C., 2004, "English dialects in the North of England: phonology", *in* Edgar W. J. Schneider, Kate Burridge, Bernd Kortmann, Rajen Mesthrie & Clive Upton (eds.), *A Handbook of Varieties of English*, Berlin, Germany, Mouton de Gruyter.

BEAL Joan C., "Enregisterment, commodification, and historical context: 'Geordie' versus 'Sheffieldish'", *American Speech*, 2009, vol. 84(2), p. 138-156.

BEAL Joan C., CORRIGAN Karen P., MEARNS Adam & MOISL Herman L., 2014, "The Diachronic Electronic Corpus of Tyneside English: Annotation practices and dissemination strategies", *in* Jacques Durand, Ulrike Gut & Gjert Kristoffernson (eds.), *Handbook of Corpus Phonology*, Oxford, Oxford University Press.

BUCHSTALLER Isabelle, KRAUSE Anne, AUER Anja *et al.*, "Levelling across the life-span? Tracing the face vowel in panel data from the North East of England", *Journal of Sociolinguistics*. 2017, vol. 21(1), p. 3-33.

CHAMBERS Jack K., 1994, *Sociolinguistic theory: linguistic variation and its social significance*, Oxford, Blackwell.

CORRIGAN Karen P., 2012, "GOAT vowel variants in the Diachronic Electronic Corpus of Tyneside English (DECTE)", *in* Terttu Nevalainen & Elizabeth C. Traugott (eds.), *The Oxford Handbook of the History of English*, Oxford, Oxford University Press.

CORRIGAN Karen P., BUCHSTALLER Isabelle, MEARNS Adam & MOISL Herman L., 2012, *Diachronic Electronic Corpus of Tyneside English [Online]*, Newcastle University, disponible sur : https://research.ncl.ac.uk/decte (consulté le 10 décembre 2022)

CORRIGAN Karen P., MEARNS Adam & MOISL Herman L., 2014, "Feature-based versus aggregate analyses of the DECTE corpus: Phonological and morphological variability in Tyneside English", *in* Benedikt Szmrecsanyi & Bernhard Wälchli (eds.) *Aggregating Dialectology, Typology, and Register Analysis: Linguistic Variation in Text and Speech*, Berlin, De Gruyter.

COUPLAND Nikolas, 1980, "Style-shifting in a Cardiff work-setting", *Language in society*, vol. 9(1), p. 1-12.

DEUMERT Andrea, 1999, "Variation and Standardisation. The Case of Afrikaans (1880-1988)", University of Cape Town.

DI FRANCO Giovanni & MARRADI Alberto, 2013, "Factor analysis and principal component analysis", Milano, FrancoAngeli.

ECKERT Penelope, 2012, "Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation", *Annual Review of Anthropology*, vol. 41, p. 87-100.

ESCOFIER Brigitte & PAGÈS Jérôme, 1984, « L'analyse factorielle multiple », *Cahiers du Bureau Universitaire de Recherche Opérationnelle*, vol. 42, p. 3-68.

GENERAL REGISTER OFFICE, 1966, *Classification of Occupations 1966*, H.M. Stationery Office.

GHYSELEN Anne-Sophie & DE VOGELAER Gunter, 2018, "Seeking Systematicity in Variation: Theoretical and Methodological Considerations on the 'Variety' Concept", *Frontiers in Psychology*, vol. 9, article 385.

GRIES Stephan, 2009, *Statistics for linguistics with R. A Practical Introduction*, Berlin, Mouton De Gruyter.

GUY Gregory R., 2013, "The cognitive coherence of sociolects: How do speakers handle multiple sociolinguistic variables?", *Journal of Pragmatics*, vol. 52, p. 63-71.

HADDICAN Bill, FOULKES Paul, HUGHES Vincent *et al.*, "Interaction of social and linguistic constraints on two vowel changes in northern England", *Language Variation and Change*, 2013, vol. 25(3), p. 371-403.

HUSSON François, JOSSE Julie & PAGÈS Jérôme, 2010, "Principal component methods – hierarchical clustering – partitional clustering: Why would we need to choose for visualizing data?", *Technical report, Agro Campus Ouest*.

HUSSON François, LE DIEN Sébastien & PAGÈS Jérôme, 2001, "Which value can be granted to sensory profiles given by consumers? Methodology and results", *Food Quality and Preference*, vol. 12, p. 291-296.

HUSSON François, LÊ Sébastien & PAGÈS Jérôme, 2011, *Exploratory multivariate analysis by example using R*, Boca Raton, CRC Press.

JANSSENS Wim, 2008, *Marketing research with SPSS,* Pearson Education.

JONES-SARGENT Val, 1983a, *The TLS phonetic transcription scheme [Online],* available at: https://research.ncl.ac.uk/decte/transscheme.htm (accessed 5 November 2022)

JONES-SARGENT Val, 1983b, *Tyne bytes: A computerised sociolinguistic study of Tyneside*, Frankfurt am Main, Peter Lang.

LABOV William, 1972, "Some principles of linguistic methodology", *Language in society*, vol. 1(1), p. 97-120.

LABOV William *et al.*, 1981, "Field methods of the project on linguistic change and variation".

LAMENTOWICZ Mariusz, LAMENTOWICZ Łukasz, VAN DER KNAAP Willem O. *et al.*, 2010, "Contrasting species—environment relationships in communities of testate amoebae, bryophytes and vascular plants along the Fen–Bog gradient", *Microbial Ecology*, vol. 59, p. 499-510.

LÊ Sébastien, JOSSE Julie and HUSSON François, 2008, "FactoMineR: An R package for multivariate analysis", *Journal of Statistical Software*, vol. 25, p. 1-18.

LEBART Ludovic, MORINEAU Alain & PIRON Marie, 1995, *Statistique exploratoire multidimensionnelle*, Paris, Dunod.

MILROY James, 1996, "Variation in /ai/ in Northern British English, with comments on Canadian raising", *University of Pennsylvania working papers in linguistics*, vol. 3, p. 213-221.

MOISL Herman L., 2015, *Cluster analysis for corpus linguistics*, Berlin, De Gruyter.

MOISL Herman L. & MAGUIRE Warren N., 2008, "Identifying the main determinants of phonetic variation in the Newcastle electronic corpus of Tyneside English", *Journal of Quantitative Linguistics*, vol. 15, p. 46-69.

NERBONNE John & WIELING Martijn, 2017, "Statistics for Aggregate Variationist Analyses", *in* Charles Boberg, John Nerbonne & Dominic Watt (eds.), *The Handbook of Dialectology*, Wiley Online Library.

PELLOWE John, NIXON Graham, STRANG Barbara & MCNEANY Vince, 1972, "A dynamic modelling of linguistic variation: The urban (Tyneside) linguistic survey", *Lingua*, vol. 30, p. 1-30.

PICARDI Michela, REDAELLI Valentina, ANTONIOTTI Paola *et al.*, 2020, "Turning and sit-to-walk measures from the instrumented Timed Up and Go test return valid and responsive measures of dynamic balance in Parkinson's disease", *Clinical Biomechanics*, vol. 80, article 105177.

SOKAL Robert R, 1966, "Numerical Taxonomy", *Scientific American*, vol. 215, p. 106-116.

STRANG Barbara, 1968, "The Tyneside Linguistic Survey", *in* Ludwig Erich Schmitt (ed.) *Verhandlungen des Zweiten Internationalen Dialektologenkongresses*, Marburg, Franz Steiner Verlag.

STRANG Barbara, 1972, "TLS 015 Tyneside Linguistic Survey (summary)", Newcastle-upon-Tyne, Newcastle University.

STRANG Barbara & PELLOWE John, 1969, "The Tyneside Linguistic Survey: General handout on settings, aims, range and methods", Unpublished report, Newcastle University.

TRUDGILL Peter, 1972, "Sex, Covert Prestige and Linguistic Change in the Urban British English of Norwich", *Language in Society*, vol. 1, p. 179-195.

WATT Dominic, 1999, "Phonetic variation in two Tyneside vowels: Levelling in lockstep", *in* John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville & Ashley C. Bailey (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences*, Berkley, University of California Press.

WATT Dominic & FOULKES Paul, 2017, "Tyneside English", *in* Raymond Hickey (ed.), *Listening to the Past: Audio Records of Accents of English*, Cambridge, Cambridge University Press. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/p14_intro.pdf

WELLS John C., 1982, *Accents of English: Volume 1*, Cambridge University Press.