RESEARCH

Inter- and intra-speaker variation in French schwa

Andrew Bayles, Aaron Kaplan and Abby Kaplan

Department of Linguistics, University of Utah, Languages and Communication Bldg, 255 S. Central Campus Dr., Rm 2300, Salt Lake City, UT 84112, US
Corresponding author: Aaron Kaplan (a.kaplan@utah.edu)

Quantitative constraint-based theories of optionality typically aim to model the frequency with which an individual speaker's grammar maps one input onto various output forms. But existing studies in this area use population-level frequency data as a proxy for the frequencies produced by an individual speaker's grammar. This practice is problematic for three reasons. First, population-level variation could result from aggregation over individuals with differing categorical behavior, not individual variation at all. Second, some theories predict that individuals with identical relative constraint rankings/weights can produce those outputs with different frequencies, while others predict uniform frequencies across speakers; these prediction cannot be tested without individual-level data. Third, if indeed individuals differ in their output frequencies, population-level averages may disguise true patterns of individual variation. This paper addresses these shortcomings through a corpus study of optional schwa deletion in French. We find that individual speakers do in fact show variable omission of schwa. We also find variation between speakers in the frequency of schwa's appearance. Taken together, these results confirm that for French schwa, formal theories are correct to ascribe variation to a single speaker's grammar, and they favor theories that decouple the means of producing variation from the account of output frequency.

1 Introduction

Optimality Theory (OT; Prince & Smolensky 1993/2004) maps any given input onto a single output candidate, but much recent research has argued that this is inadequate: often an input has multiple possible surface variants. To take a famous case, the vowel often transcribed as [ə] in French is optional in a variety of contexts (e.g. Dell 1973; Howard 1973; Selkirk 1978; Anderson 1982; Léon 1987; Tranel 1987; Van Eibergen 1991; 1992; Van Eibergen & Belrhali 1994; Côté 2001; Gess et al. 2012), so, for example, *pelouse* 'lawn' may be produced with or without the vowel in the initial syllable: [pəluz] ~ [pluz]. It is therefore incumbent upon the phonological grammar to provide both [pəluz] and [pluz] as surface forms for this word. In response to phenomena of this sort, a variety of accounts of phonological optionality has arisen. Some coopt OT's mechanism for producing crosslinguistic variation – constraint reranking – and allow multiple or variable rankings within a single language (Reynolds 1994; Nagy & Reynolds 1997; Boersma 1998; Boersma & Hayes 2001; Anttila 2006; 2007), while others, building on arguments that simple reranking is inadequate (Riggle & Wilson 2005; Nevins & Vaux 2008; Vaux 2008), develop new machinery that is responsible for producing variation within a single grammar (Coetzee 2004; 2006; Riggle & Wilson 2005; Kaplan 2011; Kimper 2011). Particular theories of optionality often go beyond the multiple-outputs threshold and supply

a means of producing one variant more often than another according to the frequencies of attestation in some corpus (e.g. Anttila 2007; Kaplan 2011).

At the heart of this research lies the claim that in some cases a speaker may produce multiple surface forms for a given lexical entry – one speaker of French can produce both [pəluz] and [pluz], e.g., a situation we will call *individual-level variation*. If, on the other hand, the variation exhibited by *pelouse* arises only across speakers – with some speakers invariably producing [pəluz] and others invariably producing [pluz], which we will call a *non-uniform population* – we're effectively dealing with different dialects, and the proper formal analysis is to posit two variation-free grammars, one for each set of speakers, just as we adopt separate grammars for different languages. This state of affairs would render theories of optionality superfluous.

Unfortunately, much research that focuses on the development of formal accounts of optionality conflates individual-level variation and non-uniform populations, and the corpora that are used to demonstrate variation combine the utterances of multiple speakers. For example, with regard to French schwa, Dell (1977) experimentally investigates the frequency of appearance of schwa at various syntactic boundaries with data from several subjects. As valuable as this study is, it does not shed light on the extent to which variation occurs within versus across speakers. Other work uses native-speaker intuitions to assess which logically possible variants are grammatical (e.g. Dell 1973; Côté 2001). While this provides useful information about forms that speakers are willing to accept, it does not tell us conclusively whether a speaker will actually produce multiple variants.

Relatedly, the frequencies-of-attestation for each variant that are used to inform analyses that aim to produce these frequencies typically derive from corpus studies. The corpus frequencies, like the examples of variation themselves, are aggregated across speakers, and this risks masking another sort of non-uniformity, one that presents non-uniform output frequencies. To return to the *pelouse* example, if one subpopulation omits [ə] 25% of the time and another does so at a 75% rate, this may average out to 50%, but a formal account of the process that predicts this 50% models no actual speaker's grammar. In building accounts of variation, then, it is essential to ensure that our models of variants' frequencies are built on an accurate understanding of the frequencies that speakers actually exhibit, and our theories must be flexible enough to accommodate non-uniform frequencies when necessary.

Pinning down the extent of individual-level variation and differences in variant frequencies between speakers can inform theories of optionality. As discussed in section 3 below, some theories of optionality necessarily predict individual-level variation, while others can be easily reinterpreted if it turns out that variation arises only across speakers. Likewise, some theories allow variants' frequencies to differ between speakers, but other theories predict the same frequencies for all speakers, at least without further elaboration. Determining whether or not (i) individual-level variation exists and (ii) output frequencies are uniform across speakers can help distinguish one theory from another and guide future development of these theories.

To this end, in this paper we add to the existing research that in one way or another – sometimes directly, sometimes incidentally – sheds light on the production patterns of individual speakers and differences between speakers (Walter 1982; Lucci 1983; Durand et al. 1987; Encrevé 1988; Hansen 1997; Eychenne 2006; 2009a; b; Hayes & Londe 2006; Poiré 2009; Pustka 2009; Bürki et al. 2011). This existing research takes various forms, from experimental to corpus-based studies (including a significant body of research that uses the corpus that we adopt for the study reported below), but the issues raised above

tend not to be the primary focus of that research, and consequently even where it can inform discussion of those questions it rarely does so in a systematic and rigorous way.[1] Our goal, therefore, is to determine the extent to which individual-level variation in the realization of [ə] in French exists and to probe differences in that variation between speakers. We present the results of a search of a corpus of French that shows that the variation in [ə]'s realization does in fact exist at the level of the individual speaker, and that speakers do not show uniform rates of schwa realization. The corpus we use, the Phonologie du Français Contemporain (PFC) corpus (http://www.projet-pfc.net/; Durand et al. 2002; 2009a), is a collection of speech from many varieties of French. It has given rise to a growing body of research (including Eychenne 2006, Pustka 2007, and the contributions to Durand et al. 2009b and Gess et al. 2012, to name just a few) and is designed to facilitate studies like ours in that it identifies schwas (or loci where schwa has been omitted) along with linguistic factors that are relevant to schwa's optionality (see section 4.1). Our strategy is this: we identify a handful of contexts in which it has been claimed that [ə] is optional, and we assess the extent to which individual speakers exhibit this optionality as recorded in the PFC corpus.

The paper is organized as follows: section 2 briefly presents background on French schwa and its optionality. Section 3 summarizes salient constraint-based theoretical approaches to optionality and examines their predictions with respect to individual-level variation and population (non-)uniformity. Section 4 presents the results of our corpus study, and section 5 evaluates the theories from section 3 in light of the corpus study. Section 6 summarizes and concludes the paper.

2 French schwa

The vowel transcribed as [ə] in French is a mid front rounded vowel whose specific realization varies across contexts and dialects; see Durand et al. (1987) and Fougeron et al. (2007), for example. It has been the subject of a large body of research, chiefly, perhaps, because it is optional in a variety of positions (e.g. Dell 1973; 1980; Howard 1973; Selkirk 1978; Anderson 1982; Léon 1987; Tranel 1987; Van Eibergen 1991; 1992; Van Eibergen & Belrhali 1994; Côté 2001, among many others). In this section we discuss three contexts for schwa that have been examined in the literature: schwa in VC___C, schwa in CC___C, and schwa in clitics. We choose these contexts because they are easily recoverable from the corpus we used for our study and are therefore central to this paper.

In the research on schwa there is often disagreement concerning the permissibility of (the absence of) [ə] in particular phonological contexts. For example, according to the Law of Three Consonants (LTC), which was first identified by Grammont (1914) and has been adopted in various forms by many researchers since, "a schwa is usually present after two consonants, while it may not be realized after a single one" (Gess et al. 2012: 5). That is, schwa may not be omitted in the context CC___C (1), thought it may be in VC___C (2). (Both sets of examples come from Dell 1980: 205.) As (2b) shows, it is not merely the prospect of a triconsonantal cluster that renders schwa obligatory. The crucial factor is schwa's location between the second and third of those consonants. (The presence of a

---

[1] Noteworthy in this respect is Eychenne (2006), who examines the behavior of schwa in a small handful of speakers in two areas of southern France (using the same corpus that serves as the basis for our study). In the course of developing an OT analysis, he compares the productions of individual speakers and often presents statistical analyses that reveal genuine differences between speakers. He does not, however, tie those results to optionality-producing formalisms in the way we attempt to do here.

consonant following schwa is crucial in its own way, too: schwa is illicit prevocalically;
Dell 1980; Côté 2001.)

(1)   a.   *j'arrive demain*
              [ʒarivdəmɛ̃], *[ʒarivdmɛ̃]
              'I arrive tomorrow'

       b.   *Jacques devrait partir*
              [ʒakdəvrɛpartir], *[ʒakdvrɛpartir]
              'Jack should leave'

(2)   a.   *j'arriverai demain*
              [ʒarivrɛdəmɛ̃], [ʒarivrɛdmɛ̃]
              'I shall arrive tomorrow'

       b.   *Henri devrait partir*
              [ãridəvrɛpartir], [ãridvrɛpartir]
              'I arrive tomorrow'

Under this view, then, schwa's omissibility is governed by the number of surrounding
consonants.

   Côté (2001), on the other hand, agrees that schwa is optional in VC___C, but she argues
that, at least in her variety of Canadian French, this optionality also extends to the
CC___C context under certain phonotactic conditions. She provides data like the follow-
ing. Omission of schwa in these forms is ruled out by the LTC, but variation is nonetheless
possible, though it may be limited to Côté's dialect.[2]

(3)   a.   *une fenêtre*
              [ynfənɛtr], [ynfnɛtr]
              'a window'

       b.   *Ester le salut*
              [ɛstɛrləsaly], [ɛstɛrlsaly]
              'Ester greets him'

Since our goal is to focus on contexts in which schwa is optional, the absence of a consen-
sus in the literature on the proper characterization of contexts that foster optionality of
schwa is not a trivial obstacle for us. Consequently, we begin by adopting the LTC's view
of optionality and examine the VC___C context. As we will see in section 4.3, particular
varieties of French adhere to the LTC to varying degrees. We therefore follow Gess et al.
(2012), a volume that presents, among other things, a region-by-region survey of schwa's
optionality from the point of view of the LTC. We probe schwa in a subset of VC___C envi-
ronments (i.e. environments that do not run afoul of the LTC) that are identified by Gess
et al. as most amenable to optionality: tokens in which schwa is in a monosyllabic clitic
(e.g. *plein de linguistes* [plɛ̃d(ə)lɛ̃gɥist] 'full of linguists'; see (12) below) or the initial syl-
lable of a polysyllabic word (e.g. *j'arriverai demain* [ʒarivrɛd(ə)mɛ̃] 'I shall arrive tomor-
row'; see (2) above). This environment, incidentally, overlaps with the contexts claimed
by Côté to foster variation, so it represents a point of agreement between the theories.

   We then turn from an LTC-based view of schwa to the one developed by Côté (2001)
in which schwa may be omitted in CC___C contexts like the one in (3). Côté argues that
schwa is generally omissible in this environment, but certain (mostly phonotactic) condi-
tions may make schwa's absence illicit. First, schwa's omission may not create a tricon-
sonantal cluster in which the middle consonant is more sonorous than its neighbors (4).
Côté argues that prevocalic /r/ is realized as an obstruent; this accounts for (4c), in which

---

[2] Unless otherwise noted, all data given here and subsequently are taken from Côté (2001).

the final member of the triconsonantal cluster is a prevocalic /r/. As an obstruent in this context, it is less sonorous than the cluster-medial [m]. (In other contexts, according to Côté, /r/ is usually a glide.)

(4)   a. *la douce mesure*
          [ladu**smə**zyr], *[ladu**smz**yr]
          'the sweet measure'

       b. *Annik le salut*
          [ani**klə**saly], *[ani**kl**saly]
          'Annik greets him'

       c. *Philippe me rasait*
          [fili**pmər**azɛ], *[fili**pmr**azɛ]
          'Philippe shaved me'

Additionally, if the middle consonant is a stop, schwa is obligatory if certain conditions are not met. In general, schwa is optional if the perceptual cues for the cluster-medial stop are well-supported in schwa's absence, though morphosyntactic considerations play a role, too. (With regard to the importance of perceptual salience, the approach draws heavily on Steriade 1999a; b; 2001.) Côté identifies three conditions that permit schwa's absence: (a) the schwa appears in a clitic, (b) the third consonant is a continuant, or (c) the first consonant is a glide,[3] a category which according to Côté includes preconsonantal /r/.

In (5) none of these conditions is met, and schwa is required. In contrast, schwa may be omitted in (6). The omitted schwa belongs to a clitic in (6a); Côté (2001: 120) characterizes data of this sort as "not unacceptable, but certainly marginal." The cluster-final consonant is a continuant in (6b).[4] In (6c), the cluster-initial segment is a glide. The second and third conditions are combined in (6d), where schwa belongs to a clitic, and the cluster-initial consonant is a glide.

(5)   a. *la douce demie*
          [ladu**sdəm**i], *[ladu**sdm**i]
          'the sweet half'

       b. *la même demande*
          [lamɛ**mdəm**ãd], *[lamɛ**mdm**ãd]
          'the same request'

(6)   a. *Alice te mentait*
          [ali**stəm**ãte], ?[ali**stm**ãte]
          'Alice lied to you'

       b. *Aline devait y aller*
          [ali**ndəv**ɛjale], [ali**ndv**ɛjale]
          'Aline had to go there'

       c. *pour demander*
          [pu**rdəm**ãde], [pu**rdm**ãde]
          'to request'

       d. *Camille te mentait*
          [kami**jtəm**ãte], [kami**jtm**ãte]
          'Camille lied to you'

---

[3] More generally, Côté argues that schwa's omission is more acceptable if the first two members of the triconsonantal cluster contrast in the feature [vocoid]. By this metric, a medial stop requires a preceding vocoid if schwa's omission is to be acceptable.

[4] Côté does not explicitly provide the transcription that includes schwa in (6b), but she strongly implies that it is acceptable. In any case, the crucial point is that schwa is not required in this example, unlike in (5).

Furthermore, depending on the morphosyntactic composition of the CC___C sequence in question, the cluster resulting from schwa's omission may be more or less acceptable. For example, as we already saw in (6a), the illicit clusters exemplified in (5) improve marginally if the interconsonantal stop belongs to a clitic. Additionally, schwa is not always required if a word boundary falls after the medial consonant (as in *casque **noir*** /ka**sk** nwar/ 'black helmet'). At derivational suffix boundaries, schwa obligatorily breaks up triconsonantal clusters (7), even in contexts like those in (6).[5]

(7)   *garderie*
        [ga**rdə**ri], *[ga**rdr**i]
        'kindergarten'

Finally, the prohibition on schwa's omission in CC___C with a medial stop is weakened if the cluster straddles a prosodic boundary (see also Lucci 1983 and Hansen 1997 for the influence of prosodic boundaries on schwa). This is illustrated by the data in (8), which comes from Côté (2001). We indicate the location of the relevant boundaries with | in the transcriptions. Each example has a larger boundary in the [kt(ə)m] sequence than the previous example, and as the size of the boundary increases, the acceptability of schwa's omission improves until, in the final example, schwa is prohibited. (The labels for the prosodic categories follow Côté 2001, with the following abbreviations: PWd = Prosodic Word; SPP = Small Phonological Phrase; MPP = Maximal Phonological Phrase; IP = Intonational Phrase.)

(8)   No boundary   *tu fais que te moucher*
                      [tyfɛ**kə**tə|**m**uʃe], [tyfɛ**kt**ə|**m**uʃe], [tyfɛ**kət**|**m**uʃe], *[tyfɛ**kt**|**m**uʃe]
                      'you only blow your nose'
        PWd   *infecte manteau*
                      [ɛ̃fɛ**kt**ə|**m**ãto], [ɛ̃fɛ**kt**|**m**ãto]
                      'stinking coat'[6]
        SPP   *insecte marron*
                      [ɛ̃sɛ**kt**ə|**m**arɔ̃], [ɛ̃sɛ**kt**|**m**arɔ̃]
                      'brown insect'
        MPP   *l'insecte mangeait*
                      [lɛ̃sɛ**kt**ə|**m**ãʒɛ], [lɛ̃sɛ**kt**|**m**ãʒɛ]
                      'the insect was eating'
        IP   *l'insecte, mets-le là*
                      *[lɛ̃sɛ**kt**ə|**m**ɛlœela], [lɛ̃sɛ**kt**|**m**ɛlœela]
                      'the insect, put it there'

The upshot is that under Côté's analysis, schwa in CC___C shows a range of behaviors from obligatory to strongly preferred to simply optional, depending on morphosyntactic and phonological factors; the LTC plays no role in her formalism.[7] The relevant factors are summarized in (9).

---

[5] An anonymous reviewer has brought to our attention evidence from the PFC corpus itself that this generalization, from Côté (2001: 85), may not be accurate. The status of schwa in this context does not affect our analysis: because the corpus does not indicate whether a suffix is present (or what kind of suffix is present), we do not investigate word-internal schwas beyond the initial syllable so as to exclude possible interference from morphology.

[6] It is our understanding that *manteau* 'coat' is masculine, and therefore the appropriate form of the adjective is *infect*. Whether this affects Côté's point here we cannot say. We assume that by replacing *manteau* with a feminine noun beginning with [m] (e.g. *mélasse* 'molasses'), the same point can be made.

[7] Hansen (1997) also investigates the effect of consonant quality on schwa, and her results are broadly consistent with Côté's analysis.

(9)

Facilitates [ə] 's Absence	Inhibits [ə] 's Absence
C_2 is not the sonority peak	C_2 is more sonorous than C_1 and C_3
C_2 stop's cues are supported	C_2 stop's cues are not supported
larger prosodic boundary	smaller/no prosodic boundary
[ə] is in a clitic	derivational suffix?

The contrast between the LTC and Côté's approach can be seen easily with the phrase *envie de te le demander* 'feel like asking you,' which contains the four schwas in bold. This phrase is often reported (e.g. Dell 1973) to have the following possible realizations.

(10) a. ãvidətələdəmãde
        b. ãvid‿tələdəmãde
        c. ãvidət‿lədəmãde
        d. ãvidətəl‿dəmãde
        e. ãvidətələd‿mãde
        f. ãvid‿təl‿dəmãde
        g. ãvid‿tələd‿mãde
        h. ãvidət‿ləd‿mãde

The set of possible outputs in (10) is consistent with the LTC. In each case, removal of one of the remaining schwas produces either a different variant on this list or an illicit variant that violates the LTC. For example, beginning with [ãvid‿tələdəmãde] (10b), removing the middle of the remaining schwas gives (10f) [ãvid‿təl‿dəmãde]. But removing the first of the three schwas gives *[ãvid‿t‿lədəmãde], an illicit form that violates the LTC by virtue of the [dtl] cluster.

    Côté argues that more outputs are possible for this phrase, and that (10) is an incomplete list. She provides the following judgments for the remaining logically possible outputs:

(11) a. ãvid‿t‿lədəmãde
        b. ??ãvidətəl‿d‿mãde
        c. *ãvidət‿l‿dəmãde
        d. ãvid‿t‿ləd‿mãde
        e. ??ãvid‿təl‿d‿mãde
        f. *ãvid‿t‿l‿dəmãde
        g. *ãvidət‿l‿d‿mãde
        h. *ãvid‿t‿l‿d‿mãde

The LTC-violating variant discussed above, [ãvid‿t‿lədəmãde], is the first one on this list. It complies with all of Côté's conditions and is thus licit: the cluster-medial stop is not more sonorous than its neighbors, and it is followed by a continuant. The omitted schwas also all come from clitics. In contrast, the ungrammatical forms in (11) all contain a triconsonantal cluster ([tld]) in which the medial consonant is the most sonorous one. The marginal forms contain a [ldm] cluster whose medial stop is not followed by a continuant or preceded by a vocoid. Perhaps the preceding liquid prevents the form from being entirely ungrammatical: a contrast in [approximant] between the first two consonants might ameliorate but not entirely eliminate the problem like a contrast in [vocoid] does (see fn. 3).

    In our investigation of schwa in CC___C, we extracted tokens from the PFC corpus that display this context, and we excluded those that do not comply with Côté's phonotactic conditions. See section 4.4.

Turning to the final context we examine, a handful of clitics are of the shape Cə,[8] and when one of these clitics appears in the context V___C, its schwa is optional:

(12)　　a.　*plein de linguistes*
　　　　　　　[plɛ̃dəlɛ̃gɥist], [plɛ̃dlɛ̃gɥist]
　　　　　　　'full of linguists'
　　　　b.　*Annie le salut*
　　　　　　　[aniləsaly], [anilsaly]
　　　　　　　'Annie greets him'

Our clitic context overlaps with our VC___C context for the LTC but not with our CC___C context. Under both the LTC and Côté's analysis, clitics in VC___C are expected to show variation. Thus we examine this context separately because it represents a clear area of agreement between those approaches. Furthermore, since our goal is to examine homogenous environments while excluding factors that prohibit variation, this narrower context allows us to focus on an even more uniform context than the previous ones.

This is a small sample of the contexts in which schwa can, must, or cannot appear. See the research cited above for a more comprehensive survey of schwa's behavior. To reiterate, we focus on the contexts presented in this section because they serve as the focus of our corpus study.

3 Theories of variation: The locus of probability

Numerous constraint-based theories of phonological optionality have been developed, and they exhibit a range of predictions concerning individual-level variation and population (non-)uniformity. Some necessarily produce individual-level variation (so discovering that such variation is unattested would threaten their viability), but others are compatible with variation at the level of the population, not the speaker. That is, they can be reinterpreted as providing exactly one output per input for each speaker, while allowing different speakers to produce different outputs. Likewise, some theories predict uniform frequencies across speakers for each variant, but others allow speakers to differ in this regard. This section surveys these theories and their predictions.

It is worth making explicit our assumption that each theory discussed below is intended to account for individual-level variation. This is rarely, if ever, made explicit in the literature supporting these theories. Perhaps the most direct reference to individual-level variation comes from Munro & Riggle (2004: 116), who observe that with respect to the optional process they examine (multiple reduplication in Pima, which they treat with a theory akin to Local Constraint Evaluation; Riggle & Wilson 2005; see below for more on this theory), "generally only memory limits the number of plurals he [their consultant] volunteers". Anttila (2007: 519) draws a distinction between variation "within" speakers and variation "across" speakers and subsequently characterizes the theories he presents in single-speaker terms (e.g. "an individual may possess multiple grammars"; Anttila 2007: 525). Otherwise, the closest acknowledgement that individual-level variation is the subject of analysis that we are aware of is, e.g., Anttila's (1997: 37) characterization of the variants he is concerned with as being in "free variation". We do not fault other researchers for failing to explicitly mention that individual-level variation is the target of their

---

[8] We take no stance on the issue of whether or not schwa is underlying (and how its status as (not) underlying may vary across contexts), but we characterize schwa in clitics as underlying for expository convenience. See Côté (2001), Côté & Morrison (2007), and Pustka (2007) for discussion of the relevant issues. An anonymous reviewer asks how epenthesis in otherwise phonotactically licit sequences would be motivated under the assumption that schwa is not underlying. We cannot explore this issue fully here, but it could be that schwa epenthesis allows for better alignment of syllable and morpheme boundaries (Kaplan to appear).

analysis; in some sense, this is a trivial point. If the data under discussion involved only variation across speakers, it would reduce to crosslinguistic variation, a phenomenon that is so central to OT that it requires none of the novel formal constructs we discuss in this section. Nonetheless, since the existence of individual-level variation is at the heart of this paper, we believe it necessary to bring this issue to the forefront.

A range of formal mechanisms has been proposed for dealing with optional phenomena in OT. In classic OT, any ranking maps an input onto a single output; this is obviously incompatible with optional phenomena, in which a single input has multiple output forms. Broadly speaking, theories of optionality cope with this in one of two ways. First, they may impose some probabilistic component on the constraint ranking so that multiple rankings are available, and each ranking provides one of the possible outputs for a given input. The second kind of theory retains a fixed constraint ranking but adopts some probabilistic system for choosing an output based on that ranking. Thus each theory includes some kind of probabilistic component, but they vary in the locus of that component. For our purposes it is most useful to examine these theories from this point of view because we will see that the second kind of theory – those in which the ranking is constant but the path from that ranking to an output is not – is most clearly compatible with the results of our corpus study.

3.1 Probabilistic selection of ranking

Perhaps the most common approach to optionality rests on the adoption of multiple constraint rankings: each ranking produces one variant, so across the available rankings, the entire range of optionality is produced. The Partial Orders (PO) model of Anttila (1997; 2007) is representative of this kind of theory. It produces variation by positing a partial constraint ranking that is resolved into a complete ranking on each evaluation. This resolution may vary across evaluations yielding different outputs for the same input across tableaux. Anttila illustrates this framework with an analysis of optional vowel coalescence in Finnish. When the second of two adjacent vowels is low, the vowels optionally coalesce:

(13)   a.   /suome-a/→ 'suomea ~ 'suomee 'Finnish (partitive)'
        b.   /ruotsi-a/ → 'ruotsia ~ 'ruotsii 'Swedish (partitive)'

Coalescence is more common when the first vowel is mid than when it is high. This is captured with the ranking *EA ≫ *IA; in the grammar at large, faithfulness is unranked with respect to these constraints, so three complete rankings are possible, shown in (14). Two trigger coalescence in [ea] sequences, but just one does so with [ia].

(14)

	Coalesce [ea]?	Coalesce [ia]?
FAITH ≫ *EA ≫ *IA	No	No
*EA ≫ FAITH ≫ *IA	Yes	No
*EA ≫ *IA ≫ FAITH	Yes	Yes

Anttila tests this partial ranking (supplemented with handful of additional constraints that capture other properties of the system) against the rate of coalescence in various contexts as revealed by a corpus of Finnish.

Does this analysis produce individual-level variation? One interpretation – perhaps the most plausible one – of the analysis is that the grammar of each speaker of Finnish contains the partial ranking so that all three total rankings in (14) are available to each speaker. This view claims that individual-level variation exists in Finnish because each speaker's grammar produces both ['suomea] and ['suomee]. Anttila argues for this position but does not show that it is necessary. It is also possible that the variation arises across speakers: we might have no individual-level variation within a non-uniform population so that

some speakers invariably produce ['suomea] and others ['suomee]. In this case, we can reinterpret the partial ranking as what we might call a "metaranking": it characterizes the population of Finnish speakers as a whole – the rankings that they share – but each speaker's grammar reflects just one of the rankings in (14) with no variation. PO, then, is viable whether or not individual-level variation exists.

The question of non-uniformity in variant frequencies under PO is harder to pin down. Under the assumption that all speakers of a particular linguistic variety have the same ranking, PO predicts uniform frequencies. Because frequencies are directly tied to the total rankings derived from a partial ranking, two speakers who have the same partial ranking are necessarily predicted to produce each variant with the same frequency. But it is easy to imagine that two speakers might have different grammars – non-identical rankings – that produce the same outputs. Since their rankings differ, they may produce variants with different frequencies. For example, by reversing the ranking of *EA and *IA in Anttila's account of Finnish coalescence, we arrive at a grammar that produces the same outputs but triggers coalescence of [ia] more often than [ea]. Frequencies could also be manipulated across speakers by changing the constraints themselves: replacing *EA with *[–LOW]A, for example, would yield a system in which [ia] is penalized by both active markedness constraints while [ea] is penalized by just one. However, we are not aware of any explicit proposal for accommodating non-uniform frequencies in these ways; the working assumption seems to be that all speakers have the same partial rankings. Which interpretation of PO is the correct one is an empirical question, and in that respect our study can be seen as a means of adjudicating between these options. We return to this issue in section 5.

What is clear, though, is that PO permits a very particular kind of non-uniform frequencies. With variants' frequencies dictated by the subset of the factorial typology a speaker has access to, any particular set of constraints permits only a discrete set of possible frequencies. To illustrate with the Finnish analysis, if there are no fixed rankings among the three constraints, there are $3! = 6$ possible rankings; for any candidate, it will win in n of these rankings, so since each ranking has a $\frac{1}{6}$ chance to emerge, each candidate's probability is $n \cdot \frac{1}{6}$. Only multiples of $\frac{1}{6}$ are possible frequencies.

Suppose we impose a ranking between two of the constraints – say, *EA ≫ *IA as in Anttila's analysis. Half of the factorial typology is now eliminated (the remaining rankings are given in (15)), and each candidate's probability is $n \cdot \frac{1}{3}$. This is just a subset of the original multiples of $\frac{1}{6}$. Imposing a second ranking eliminates either one or two of the rankings in (15). FAITH ≫ *EA eliminates all but the first ranking, and *IA ≫ FAITH eliminates all but the last ranking. *EA ≫ FAITH eliminates just the first ranking, and FAITH ≫ *IA eliminates just the last ranking. If one ranking remains, each candidate's probability is just n; if two remain, each candidate's probability is $n \cdot \frac{1}{2}$. Once again, the available frequencies are a subset of the frequencies made available by the original factorial typology. Given just three constraints, then, at most the frequencies $0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}$, and 1 are permitted.

(15)   a. FAITH ≫ *EA ≫ *IA
          b. *EA ≫ FAITH ≫ *IA
          c. *EA ≫ *IA ≫ FAITH

More generally, in a PO analysis using c constraints, each candidate's frequency is some multiple of $\frac{1}{c!}$. This predicts that speakers' output frequencies will be clustered at discrete intervals around the multiples of $\frac{1}{c!}$. Obviously more gradient-seeming frequencies – and more precise frequency predictions – can be generated by introducing more constraints,

but for a theory in which frequencies are supposed to be natural consequences of the constraint ranking, this strategy seems counter to the spirit of the formalism.

In our study we do not have the data to say for sure whether PO's discrete-intervals prediction is borne out. (Such an evaluation would also require a full PO analysis of French schwa, but we are aware of no such proposal.) PO's predictions in this regard are quite different from other theories, however, as we will see.

PO is just one of a number of theories that rely on variation within the constraint ranking. Whereas PO exploits a partial ranking, others list the rankings directly (Anttila 2006) or posit floating blocks of constraints that can appear at different points in the ranking on different evaluations (Reynolds 1994; Nagy & Reynolds 1997). With respect to the kinds of variation these theories predict, they all behave similarly to PO because they all rely on the same basic mechanisms to derive variation and output frequencies.

Stochastic OT (S-OT; Boersma 1998; Boersma & Hayes 2001), like PO, attributes variation to the availability of multiple constraint rankings, but it implements that approach via a continuous and stochastic conception of rankings. On each evaluation, noise is added to the constraint ranking so that A ≫ B may become B ≫ A, and variation results. The variation produced by this mechanism is inherently individual-level variation: if two constraints are ranked sufficiently close together so that the added noise makes their ranking variable, there is no way to ensure that for each speaker the ranking settles out the same way on every evaluation. That is, unlike the situation with PO, we see no way of reinterpreting S-OT's formalism so that it provides variation across but not within speakers. Furthermore, by changing the ranking probabilities for constraints across speakers (without changing the relative dominance relationships), we can produce the same range of variation for different speakers but different variant frequencies.

Serial Variation (SV; Kimper 2011) combines PO with Harmonic Serialism, a theory described by Prince & Smolensky (1993/2004) in which outputs are produced incrementally via multiple passes through the grammar. If a partial ranking is resolved differently on different iterations, an optional process may be triggered at one locus on one iteration but not at a different locus on a subsequent iteration. Since SV allows the ranking to change within a derivation, it necessarily predicts that a single speaker will be able to produce multiple surface forms for a single input. But its mechanism for modeling variant frequencies is the same as PO's (i.e. one variant is more likely than another if there are more resolutions of the partial ranking that produce it), so its predictions regarding inter-speaker variation are the same as PO's: it produces non-uniform frequencies if the details of the partial ranking are allowed to vary across speakers.

Finally, Local Constraint Evaluation (LCE; Riggle & Wilson 2005) decomposes each constraint into position-specific instantiations: *SCHWA@1 assigns a violation when position 1 in a candidate contains a schwa, but it assigns no violations for schwas in other positions. Adopting a variable ranking between the parent constraints *SCHWA and MAX means that the position-specific MAX constraints can be freely ranked with the position-specific *SCHWA constraints. While one evaluation might operate under *SCHWA@1, MAX@2 ≫ MAX@1, *SCHWA@2 and thereby delete a schwa in position 1 but not position 2, we could arrive at the opposite ranking on another evaluation, leading to deletion of just the schwa in position 2.

Whether LCE requires individual-level variation is not clear. If position-specific constraints are projected on each evaluation according to the number of positions in the form at hand, then it is difficult to see how, say, the interleaving of the *SCHWA and MAX constraints could be held constant across evaluations (and thus the theory would necessarily predict individual-level variation). But if position-specific constraints are full-fledged

members of CON, then the theory reduces to PO in the relevant respects, and it can be reinterpreted as providing a non-uniform population with no individual-level variation if necessary. Under either view, however, non-uniform frequencies result from manipulation of the constraints responsible for variation, just like PO.

3.2 Probabilistic selection of output

The second kind of theory produces optionality by introducing some way of making multiple outputs available under a single ranking. For example, the Rank-Ordered Model of Eval (ROE; Coetzee 2004; 2006) posits that all candidates that survive to a designated point in an evaluation are possible outputs. Those that are ruled out before this cut-off line are eliminated from consideration as normal. This theory produces multiple outputs with a single constraint ranking, and since there is no way to designate one of those candidates as the winner across all evaluations for a particular input and speaker, it necessarily predicts that a given speaker will have access to multiple surface forms. ROE is intentionally vague about frequency predictions, claiming only that more harmonic outputs should be more frequent than less harmonic ones. More specific frequency predictions are left to other factors (e.g. sociolinguistic considerations). This leaves plenty of room for these other factors to trigger different variant frequencies for different speakers.

Markedness Suppression (MS; Kaplan 2011) produces variation by allowing Eval to discard violation marks at random for markedness constraints that receive a special designation on a language-particular basis. Kaplan applies this framework to variation in French schwa. The analysis designates *SCHWA as *suppressible*, meaning that violations it assigns can be discarded; this is indicated by ⊙*SCHWA. The tableau in (16) shows how this framework produces the variation seen in the phrase *envie de te le demander* 'feel like asking you' (see (10) and (11) in section 2), using the LTC-based grammaticality judgments for simplicity. Each of the indicated winners is a possible output if its violations of ⊙*SCHWA are removed,[9] but candidate (i) cannot win because it is eliminated by higher-ranking constraints against the [dtldm] cluster.

(16)	/ãvidətələdəmãde/	PHONOTACTICS	⊙*SCHWA	MAX
☞	a. ãvidətələdəmãde		****	
☞	b. ãvidtələdəmãde		***	*
☞	c. ãvidətlədəmãde		***	*
☞	d. ãvidətəldəmãde		***	*
☞	e. ãvidətələdmãde		***	*
☞	f. ãvidtəldəmãde		**	**
☞	g. ãvidtələdmãde		**	**
☞	h. ãvidətlədmãde		**	**
	i. ãvidtldmãde	*!		****

This theory models only individual-level variation. There is no mechanism by which we can always eliminate the ⊙*SCHWA violations for only, say, [ãvidətlədəmãde] for one speaker and [ãvidtələdmãde] for another. Variation arises within a single ranking by manipulating violation marks in a way that cannot be held constant for a single speaker.

---

[9] More precisely, each of these candidates wins when the violations of ⊙*SCHWA that are suppressed on a given evaluation leave a pattern of violation marks that renders that candidate optimal.

On the other hand, MS allows variants to be produced at different rates across speakers. The likelihood of suppression for any violation mark is specified by a probability p; under $p = .5$, e.g., each violation mark has a 50% chance of suppression. By specifying different values of p for each speaker, we can produce different output frequencies across speakers.

Finally, Maximum Entropy grammars (MaxEnt; Goldwater & Johnson 2003; Jäger & Rosenbach 2006; Jäger 2007) assign a probability to each candidate based on its performance on the constraint weighting: more harmonic candidates have a greater probability. (See especially Jesney 2007 for a study of MaxEnt's applicability to optionality.) Under MaxEnt, violations are interpreted numerically, as in Harmonic Grammar (Legendre et al. 1990; Smolensky & Legendre 2006), whereby the number of violations of a constraint is multiplied by that constraint's weight so that candidates that violate higher-weighted constraints receive greater penalties. Since every candidate in MaxEnt has some non-zero probability of being chosen as the output, the theory produces individual-level variation. Non-uniform frequencies can result from manipulating constraint weights so that candidates' probabilities vary across speakers.

3.3 Summary

The theoretical landscape surveyed here is summarized in (17). Clearly, the presence or absence of individual-level variation and non-uniform populations/frequencies in French is important for more than what it tells us about French specifically. Determining whether individual speakers show variation in the realization of schwa is an important step in assessing the viability of theories that require this sort of variation. If variation in French schwa (or Finnish coalescence, for that matter) arises only across speakers, these frameworks are unsuitable. Even worse, if all optional processes behave similarly, with variation arising only across speakers, these theories model a nonexistent phenomenon. Likewise for non-uniform variant frequencies: some theories predict such non-uniformity, and others do so only under certain interpretations (interpretations that have not been explicitly developed in the literature, moreover), so determining whether or not frequencies are uniform across speakers can provide empirical support for or against particular theories.

(17)

	PO	S-OT	SV	LCE	ROE	MS	MaxEnt
Requires individual-level variation	✕	✓	✓	?	✓	✓	✓
Permits non-uniform frequencies	?	✓	?	?	✓	✓	✓

It is worth noting that a number of these theories – SV, LCE, and MS, especially – were conceived to address what Riggle & Wilson (2005) call *local optionality*, a particular kind of optionality that PO and related theories are seen as incapable of producing (but see Kaplan to appear for an argument that PO can produce local optionality under the right circumstances). Under local optionality, when a form contains multiple loci of variation, those loci behave independently of each other. For example, in *envie de te le demander,* there is no requirement that the optional schwa deletion/insertion process apply in the same way to all four loci. This is unexpected under PO: to take all four schwas as underlying for purposes of illustration, a variable ranking between MAX and *SCHWA produces either no deletion (MAX ≫ *SCHWA) or the most deletion allowed by the LTC or other phonotactic constraints (*SCHWA ≫ MAX). Mechanics introduced by other theories, such as LCE's position-specific constraints, are designed to grant OT more flexibility.

Thus these theories grow out of the effort to construct a formalism that provides all and only the possible outputs in optional phenomena, and much of the literature cited above with respect to theses theories is concerned with measuring formalisms against this metric. That is clearly important work, and our goal here is to add to that strand of research other criteria by which theories of optionality can be evaluated: do they place variation in the right place (individuals versus populations), and do they provide the right level of flexibility with respect to output frequencies?

In what follows, we present the results of a corpus study aimed at assessing the nature of the variation of French schwa by searching for evidence of individual-level variation. But it is not sufficient to simply show that a single speaker produces schwa in some utterances but not others. Many factors have been argued to affect schwa's realization, and if schwa is truly optional (in some circumstances), these factors must be weeded out. For example, as we've seen, certain phonological contexts require schwa and others forbid it; lumping these contexts together to show that schwa appears only some of the time would be misleading. Discourse type affects schwa's appearance (Hansen 1994), so it would be equally dangerous to show that schwa is variable in a corpus that includes a variety of discourse types: perhaps all the tokens with schwa are from formal utterances and all the tokens without it are from informal ones. In this case, schwa is not necessarily variable; instead, we've uncovered evidence for two different grammars, one which requires schwa and another which bans it. Fortunately, the PFC corpus allows us to control for these factors, at least to a degree. The phonological context for each token is provided by the corpus, and the corpus is composed of a handful of discourse contexts according to which tokens may be segregated.

4 Corpus analysis
4.1 *Overview of the PFC corpus*

An investigation of schwa's variation must carefully navigate phonological, discourse, and dialectal factors. As we've seen, this vowel's realization may be conditioned by these considerations, so simply determining that [ə] is realized in a subset of a particular speaker's tokens in which it could appear does not show that schwa is in fact optional for that speaker: what appears to be variation may actually be consistent realization in one context and consistent omission in another. Likewise, in comparing frequencies of schwa omission across speakers, it is necessary to control for dialect: non-uniform frequencies between speakers may reflect different dialects, not different rates of schwa omission within a single dialect.

The PFC corpus (http://www.projet-pfc.net/; Durand et al. 2002; 2009a) allows a degree of control over all of these factors. It is an audio corpus of French, with accompanying transcriptions; in a portion of the corpus, potential and actual sites for schwa are coded for many of the factors discussed above. It is this subset of the corpus that we analyze in this paper.

The corpus represents speakers from Francophone areas around the world and includes detailed demographic information about each speaker. To our knowledge, the extant literature on schwa's variability is most robust for Parisian and Canadian varieties of French (though many other varieties have been studied; see especially Durand et al. 2009b and Gess et al. 2012). For this reason, we restricted our analysis to the 47 speakers identified as residents of either Île de France (the region including Paris) or Canada. Beyond place of residence, the corpus has fields for potentially relevant demographic information for each speaker, such as age and level of education. Unfortunately, not all information is supplied for each speaker, and these additional demographic categories are provided for too few speakers and tokens to allow us to draw reliable conclusions about them. Consequently, we limit our analysis of individual-level factors to place of residence.

The corpus provides a total of 23828 tokens of potential schwas – positions in a word in which a schwa can appear if the word or morpheme surfaces in the proper context – for the Île de France and Canadian speakers; of these, 22% were produced with [ə], 78% were produced without [ə], and 117 are coded as unclear. These 'unclear' productions were excluded from further analysis.

Our choice of contexts to investigate is constrained by the potential for matching contexts coded in the corpus with contexts described in the literature. In practice, this means we are able to probe just the contexts identified in section 2. Each token is coded for the following factors which we included in our analysis.

- Discourse type (Discourse): read text, guided conversation, and free conversation. All three were included in our analyses. The corpus also includes some readings of word lists, but tokens of this type were not available for the speakers we analyzed.
- Word size and position in the word (Word): monosyllabic word, first syllable of a polysyllabic word, interior (i.e. neither initial nor final) syllable of a polysyllabic word, and last syllable of a polysyllabic word. An additional code for 'metathesis' indicates situations in which a schwa precedes a consonant it would normally follow; there were 5 such tokens in our dataset, which we excluded from further analysis. For the LTC-based analysis of VC___C, we also excluded the interior syllable of a polysyllabic word and last syllable of a polysyllabic word, positions which do not show variation in the dialects under consideration (see section 4.3). For the CC___C context, we examined only tokens from monosyllabic words or the first syllable of a polysyllabic word. The other contexts were excluded to avoid influence from morphology (see the discussion above (7)) and prosodic boundaries (see (8)). For our investigation of clitics, we limited our study to the monosyllabic words: all of the monosyllabic words containing schwa are clitics, and all clitics containing schwa are monosyllabic.

Furthermore, each token is coded for the following additional factors. We used these to select the tokens to be analyzed.

- Left-hand context (Left): V(#)C___, C(#)C___, start of an intonation group, uncertain schwa to the left, and simplified consonant group. These last two codes indicate contexts in which the schwa was preceded by another potential schwa whose presence was unclear to the coder or by a cluster in which not all of the consonants were realized; we excluded these 500 tokens from our analysis to ensure a homogenous dataset. We also excluded the 8 tokens where Left was coded as the beginning of an intonational phrase but Word was coded as an interior or final syllable of a polysyllabic word, as the interpretation of this context was unclear. For our LTC-based analysis and our analysis of clitics, we included only the V(#)C___ tokens. For the CC___C context, we used only the C(#)C___ tokens. In all three cases, tokens at the start of an intonational group were excluded to ensure a more homogenous dataset and avoid the complications introduced by prosodic boundaries (see (8)).
- Right-hand context (Right): ___V, ___C, strong intonational boundary or end of utterance, and weak tonic boundary. We used only the ___C tokens in our analyses: schwa is generally prohibited prevocalically (Dell 1980; Côté 2001), and we sought to exclude the effects of prosodic boundaries from our analysis;

furthermore, in this case there were insufficient tokens that carried codes for boundaries. We also omitted 1 token coded as 'NA' for Right.

To summarize, we identified tokens for our three contexts by combining the factors listed above in appropriate ways. For the LTC-based analysis of VC___C, we used schwas in monosyllabic words and in initial syllables of polysyllabic words that were preceded by VC and followed by C. For the CC___C context, we used schwas in monosyllabic words and the first syllable of a polysyllabic word where the schwa was preceded by C#C and followed by C.[10] For clitics, we used only monosyllabic words with a schwa preceded by V#C and followed by C. In all VC___ contexts, there is the possibility that the initial V is itself a schwa; likewise, in CC___ contexts, it is possible that schwa has been omitted between the two consonants (i.e. the CC sequence alternates with CəC). We did not exclude any such tokens from our study because the code for Left indicates whether a vowel (including schwa) is present and therefore whether the token in question belongs in the VC___ category or the CC___ category. Neither the LTC nor Côté's analysis treats VC___ where V = [ə] differently from VC___ where V is another vowel, so we also chose not to distinguish them. Likewise for CC___ in which a schwa between the consonants is possible and CC___ where it is not.

With certain tokens excluded as described in the foregoing discussion, the final dataset contained 23197 tokens, and speakers contributed a median of 453 tokens each. The smallest number of tokens contributed by any speaker was 117, and the largest number was 956.

4.2 Analysis

In what follows, we analyze the occurrence of [ə] with respect to the contexts discussed above. For each phonological context, we ask two questions. The first question is whether there is in fact individual-level variation in the use of [ə]. In each case, the answer is yes: most or all of the speakers in the corpus produce [ə] some, but not all, of the time in a given phonological context.

The second question we consider is whether all speakers produce [ə] in that context with the same frequency. To address this question, we adopt the simplifying assumption that each speaker s_i has a fixed probability p_i of producing [ə] in that context. The question, then, is whether it is likely that all 47 speakers have the *same* probability p – in other words, whether all speakers have the same grammar (for that context). Since the dataset represents a random sample of speakers' productions, not every speaker will be observed to produce [ə] at a rate of exactly p_i. Thus, our question is whether the observed variation in speakers' rates of [ə] production is consistent with the hypothesis that all speakers have the same probability of producing [ə] in a given utterance.

To test this question, we fit a series of mixed-effects logistic models predicting the occurrence of [ə] in each phonological context. Every model predicts the outcome Schwa (1 for yes, 0 for no) from the fixed effect Word (i.e. word size and position) and the random effect Discourse.[11] Thus, our base model for each context is a random-intercept model with the structure

(18)　　$y_i = \alpha_{j[i]} + \beta x_i + \varepsilon_i,$

where $\alpha_{j[i]}$ is the baseline probability of realizing [ə] in discourse type j, β is the effect of Word, and ε_i is the error term.

---

[10] The presence of a word boundary in this context was invariant: the monosyllabic words under discussion are all of the shape CV, and the PFC corpus excludes CC-initial polysyllabic words from the set of tokens coded for schwa.

[11] We modeled Discourse as a random effect because we consider the discourse types represented in the corpus – read speech, guided conversation, and free conversation – to represent a sample of the larger set of possible discourse types.

For each context, our primary question is whether all speakers have the same probability of producing [ə]. To show that the population is indeed non-uniform, we must establish two things: (1) that there are significant differences among speakers in their frequency of [ə], and (2) that these differences are not just dialect-level differences.

To explore the first question, we compare our baseline model to a model with an additional random intercept for Speaker. If the two models do not differ significantly in how well they fit the data, as measured by an analysis of variance (ANOVA), then we can conclude that speakers do not differ in their overall probability of producing [ə]. But if the model with a random effect of Speaker is significantly better than the one without, then we can conclude that speakers do in fact differ.

To explore the second question, we use the factor City as a crude approximation of dialect-level variation, and treat the factors City and Speaker as competitors. We build a superset model that includes both City and Speaker as random intercepts; if this model is significantly better than a model with City only – but *not* significantly better than a model with Speaker only – then we can conclude that City does not explain any variation in the data beyond what is explained by Speaker, and therefore that the variation we see is truly between individual speakers.

Because we build a separate statistical model for each phonological context, we are unable to say anything about more general effects of the phonological or morphosyntactic features that define these contexts. We adopt this procedure because we are interested, not in the factors that influence the realization of [ə] by themselves, but rather in what the variability of [ə] looks like once known sources of non-optionality have been factored out. For example, all three subsets of the data include clitics; however, we are unable to say anything about how clitics in general affect the realization of [ə] (or how they interact with other factors). Such a question is beside the point: we are not in fact interested in clitics *per se*; we want to zero in on environments that are known in the literature to foster variation, and the environments we analyze all happen to include clitics.

4.3 The Law of Three Consonants

In this section we present analyses that assume that the Law of Three Consonants accurately captures the environments that lead to schwa's variability. In particular, we follow Gess et al.'s (2012) implementation of the LTC: "a schwa is usually present after two consonants, while it may not be realized after a single one" (Gess et al. 2012: 5). According to this view, schwa is required (or at least strongly favored) in the context CC___C, whereas it is optional in the context VC___C. Thus in this section we include data like that in (2) (e.g. *j'arriverai demain* 'I shall arrive tomorrow' [ʒarivrɛdəmɛ̃] ~ [ʒarivrɛdmɛ̃]) but not data like that in (3) (e.g. *une fenêtre* 'a window' [ynfənɛtr] ~ [ynfnɛtr]) because only in the former is schwa predicted by the LTC to be optional.

Gess et al.'s simple statement of the LTC requires modifications for particular dialects. For this reason we focus on the speakers from Île de France and Quebec from the corpus and present separate analyses of each dialect. We take as our guides for these dialects Hansen (2012) and Côté (2012), who describe the behavior of schwa in Parisian and Quebec French, respectively, using the PFC corpus.

4.3.1 Parisian French

Hansen (2012) finds that while the LTC is generally a good guide to Parisian speakers' behavior, certain contexts require special mention. First, she finds that [ə] almost never appears in word-medial VC___C tokens (i.e. schwa does not appear after a single consonant in non-word-initial, non-word-final syllables), and that it is extremely rare in word-final syllables in this context. Our findings in the PFC corpus for the Île de France region are similar: only

4.71% (19 out of 403 tokens) of word-medial tokens, and 1.11% (53 out of 4790 tokens) of word-final tokens, realize [ə] in this context. Thus, we omit word-medial and -final syllables from further analysis because they do not appear to foster robust variation.
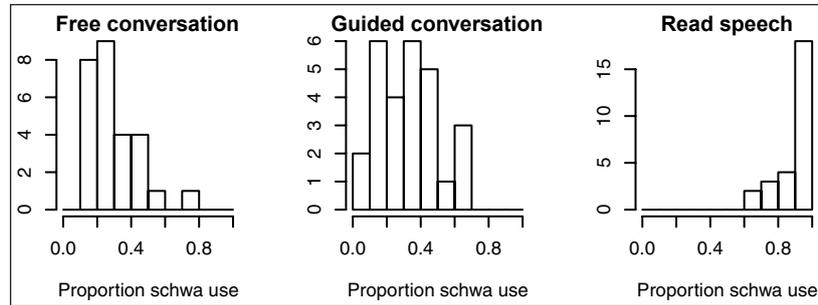
Figure 1: Histograms of individual Île de France speakers' frequency of word-initial and clitic [ə], broken down by discourse type.

The final dataset comprises 2359 tokens, of which 1125 (47.7%) are realized with [ə]. As shown in Figure 1, there is clear individual-level variation: in free and guided conversation, the vast majority of speakers neither invariably produce nor omit [ə]. (In read speech, unsurprisingly, [ə] is much more frequent, and many speakers approach categorical behavior.)

Turning to our second question – whether speakers have different probabilities of producing [ə] – we find that a model with a random effect of Speaker improves significantly over a baseline model without it ($\chi^2(3, 4) = 49.3$, $p = 2.18e - 12$). A model with a random effect of City does not improve on the baseline model ($\chi^2(3, 4) = 2.34e - 10$, $p = 1$); unsurprisingly, a model with both Speaker and City improves on the model with City only ($\chi^2(4, 5) = 49.3$, $p = 2.18e - 12$), but not on the model with Speaker only ($\chi^2(4, 5) = 0$, $p = 1$). We conclude that in this context, the population is genuinely non-uniform in terms of variant frequencies, and this non-uniformity cannot be reduced to dialect-level variation.

	Estimate	St. Error	z-value	Pr(>|z|)
(Intercept)	0.11	0.88	0.12	0.90
Word: 1st σ of polysyllabic	0.26	0.13	1.94	0.05

Table 1: Fixed effects of the final model of VC___C schwa in Île de France.

Table 1 shows the fixed effects of the final model. The effect of Word is marginally significant; the positive sign of the estimate for this factor shows that [ə] is somewhat more likely to appear in the first syllable of a polysyllabic word than it is in the baseline condition (monosyllabic words, which are all clitics). This is consistent with claims in the existing literature (Côté 2001; Hansen 2012).

4.3.2 Quebec French

Côté's (2012) study shows that Quebec French is roughly consistent with the LTC. Like Hansen, Côté finds that [ə] does not appear word-medially in the context VC___C or word-finally, a claim also supported by our data: among speakers from Quebec, [ə] is realized in 1.89% (2 out of 106 tokens) of word-medial tokens, and 1.52% (15 out of 986 tokens) of word-final tokens. With so little variation in these positions, we excluded them from our analysis. In monosyllabic words and in word-initial syllables, Côté reports variation in both VC___C and CC___C. Variation in the latter context is inconsistent with the LTC, so we excluded it, too. (These tokens are analyzed in the following section.) We are left with monosyllables and the initial syllable of polysyllabic words in the context VC___C.
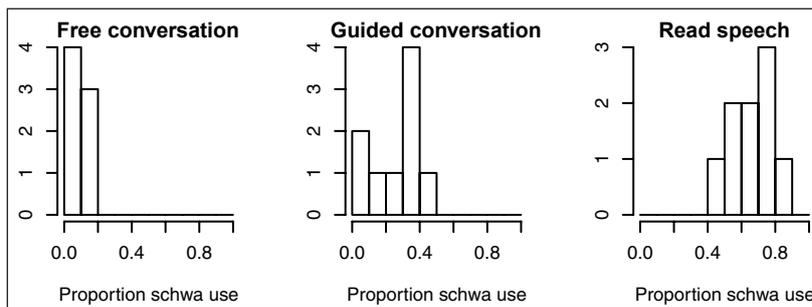
Figure 2: Histograms of individual Quebecois speakers' frequency of word-initial and clitic [ə], broken down by discourse type.

As shown in Figure 2, the final dataset clearly shows non-categorical behavior by individual speakers, particularly in guided conversation and read speech. However, we find no evidence for non-uniform frequencies across speakers in Quebec French: a model with a random effect of Speaker does not significantly improve on a model without it ($\chi^2(3, 4) = 2.4, p = 0.121$). (Because the Quebec data includes speakers from only one city, we are not able to test for a random effect of City.) However, it is possible that this null result is due to inadequate power: we have only 9 speakers from Quebec, as compared to the 27 speakers from Île de France.

4.4 Schwa in CC___C: Côté (2001)

As discussed above, the context CC___C is a much more nuanced environment under the analysis of Côté (2001) than it is under the LTC. This is not a homogenous environment for Côté: sometimes schwa is obligatory, and some-times it is optional. The summary table of the factors influencing schwa under Côté's approach is repeated below.

(19)

Facilitates [ə]'s Absence	**Inhibits [ə]'s Absence**
C_2 is not the sonority peak	C_2 is more sonorous than C_1 and C_3
C_2 stop's cues are supported	C_2 stop's cues are not supported
larger prosodic boundary	smaller/no prosodic boundary
[ə] is in a clitic	derivational suffix?

In this section we focus on the CC___ C tokens (like *une fenêtre* 'a window' [ynfənɛtr] ~ [ynfnɛtr]) that were excluded in the previous section, and we do not include any VC___C tokens. To narrow the focus to those CC___C environments that Côté identifies as exhibiting optional schwa, we consider here only CC___C involving schwa in the first syllable of a polysyllabic word or in a clitic. This eliminates contexts involving derivational suffixes, which Côté claims render schwa obligatory (see (7), repeated below).

(20)   *garderie*
        [gardəri], *[gardri]
        'kindergarten'

Also, while the PFC corpus provides information concerning the presence of prosodic boundaries, there are not enough such tokens in our dataset to support an analysis that includes boundary information; we therefore exclude tokens coded as involving prosodic boundaries from the analysis presented here.

Finally, to deal with the segmental contexts that preclude variation, the first author (a near-native speaker of Canadian French) transcribed the consonants in all 1232 CC___C tokens for six of the eleven Center-of-Paris speakers and all 9 Quebecois speakers. (To ensure that we had a sufficient number of tokens to analyze, we selected the six Parisian

speakers with the largest numbers of tokens.) To check for reliability, another near-native speaker (not affiliated with this study) transcribed a random sample of 50 of these tokens; agreement between the two coders was 92%.

Recall that Côté argues that schwa's absence cannot yield a triconsonantal cluster in which the middle consonant is more sonorous than both of its neighbors.[12] Relevant data are repeated in (21).

(21)　　a.　*la douce mesure*
　　　　　　[ladu**smə**zyr], *[ladu**sm**zyr]
　　　　　　'the sweet measure'
　　　　b.　*Annik le salut*
　　　　　　[ani**klə**saly], *[ani**kl**saly]
　　　　　　'Annik greets him'
　　　　c.　*Philippe me rasait*
　　　　　　[fili**pmə**razɛ], *[fili**pmr**azɛ]
　　　　　　'Philippe shaved me'

We discarded all 60 tokens with this property, of which 59 were realized with schwa, suggesting that schwa is indeed obligatory (or nearly so) in this environment. Furthermore, Côté claims that schwa's absence cannot yield a triconsonantal cluster in which the middle consonant is a stop unless one of the following conditions holds: (a) the schwa appears in a clitic (i.e., our monosyllabic words; (22a)), (b) the following consonant is a continuant (22b), or (c) the preceding consonant is a glide (22c), a category which according to Côté includes preconsonantal /r/.

(22)　　a.　*Alice te mentait*
　　　　　　[ali**stə**m̃ate], ?[ali**stm**̃ate]
　　　　　　'Alice lied to you'
　　　　b.　*Aline devait y aller*
　　　　　　[ali**ndə**vɛjale], [ali**ndv**ɛjale]
　　　　　　'Aline had to go there'
　　　　c.　*pour demander*
　　　　　　[pu**rdə**m̃ade], [pu**rdm**̃ade]
　　　　　　'to request'

We excluded 10 tokens on these grounds. (Of these tokens, 2 lacked the supposedly obligatory schwa; although we lack sufficient data for analysis, it is interesting to note that in both cases schwa was absent from the initial syllable of the word *petite* "small (fem.)".)

For these analyses, we accounted for the variability in the realization of /r/ in French (see section 2 as well as Côté 2001 and references therein) by treating prevocalic /r/ as an obstruent and /r/ in other positions as a glide.

As shown in the histograms in Figure 3, there is indeed individual-level variation in the realization of schwa in the remaining tokens, at least in the Free and Guided Conversation discourse types, confirming Côté's claim that schwa is optional in this subset of CC___C tokens.

Turning now to the question of whether the population of speakers is uniform, we find that adding a random effect of Speaker significantly improves the baseline model ($\chi^2(3, 4) = 9.31, p = 0.00227$). The random effect of City does not improve the baseline model ($\chi^2(3, 4) = 0, p = 1$); moreover, a model with both Speaker and City improves

---

[12] For purposes of the generalization under discussion, Côté adopts the following sonority scale, which we also adopt: Obstruents < Nasals < Liquids < Glides.

over a model with just City ($\chi^2(4, 5) = 0$, $p = 0.00227$), but not over a model with just Speaker ($\chi^2(4, 5) = 9.31$, $p = 1$). Thus, there are in fact differences among speakers in how often they produced schwa in the context CC___C, and these differences cannot be attributed to dialect-level variation. Fixed effects of the final model are shown in Table 2.

	Estimate	St. Error	*z*-value	Pr(>|z|)	
(Intercept)	2.77	0.87	3.18	0.00	*
Word: 1st σ of polysyll.	0.41	0.76	0.54	0.59	

Table 2: Fixed effects of the final model of CC___C schwa.

To summarize, once tokens of CC___C that possess consonantal properties that preclude schwa's omission are excluded, we find that schwa is indeed optional in this environment. Not only does variation within speakers exist (outside of read speech, at least), but we find variation across speakers.
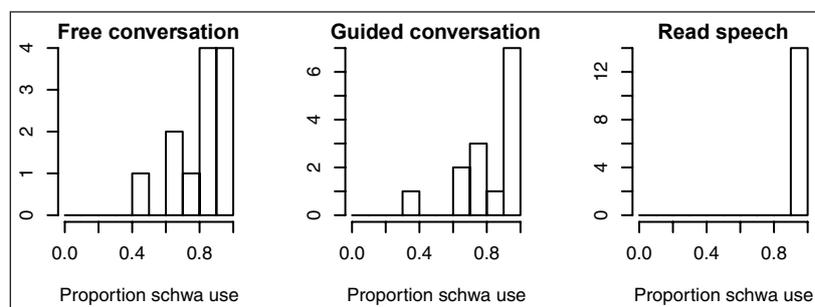
Figure 3: Proportion schwa use in the context CC___C by speaker.

That is, speakers do not behave homogenously with respect to the rate of schwa's appearance in this context, and these differences cannot be attributed to dialect-level variation.

4.5 Clitics

In this section we probe schwa's behavior in clitics more thoroughly. While consonantal properties potentially interfere with schwa's optionality in CC___C, Côté (2001) states that schwa in clitics is always optional with just one preceding consonant, an environment in which the LTC also predicts variation. Her examples, including those in (12) above, are limited to cases in which the clitic is not utterance-initial but is preceded by another (vowel-final) word; utterance-initial clitics do not behave quite so uniformly (Malécot 1976). We therefore restrict ourselves to the V#C___C environment. The data includes all of these tokens from Canada (and not just Quebec, unlike the previous analyses) and Île de France, so while there is overlap between this context and the tokens we used for the LTC-based analysis from section 4.3, the current tokens are not a subset of the data from that section.

As shown in the histograms in Figure 4, speakers exhibit a wide range of frequency of [ə] in clitics in this environment; there are a few speakers with categorical behavior at either end of the scale (i.e., speakers who never or always produce [ə] in this context in at least some discourse type), and many more speakers who fall somewhere in the middle. Because the data is broken down by discourse type, we see that there is substantial variation even within a single discourse context; thus, the observed variation in the use of [ə] in clitics is not an artefact of pooling data across categorical behavior in a variety of discourse types.

Our baseline models for the clitic dataset do not include a fixed effect for Word, since all tokens have the same value for this factor (monosyllabic words). A random effect of

Speaker significantly improves on a model without it (χ^2(2, 3) = 60.7, p = 6.56$e-15$); thus, there are real differences among speakers in the frequency of [ə] in this environment.

A model with a random effect of City is superior to the model with neither random effect (χ^2(2, 3) = 40.2, p = 2.32$e-10$). The model with *both* random effects is superior to a model with only a random effect of Speaker (χ^2(3, 3) = 7.18, p = 0.00736), and also superior to a model with only a random effect of City (χ^2(3, 4) = 27.7, p = 1.4$e-07$). Thus, for [ə] in clitics, we find evidence of non-uniformity at *both* the individual level *and* the dialect level. A population-level measure of the frequency of [ə] in clitics, even if drawn from a single dialect area, would therefore fail to capture the true range of grammars possessed by French speakers. Fixed effects of the final model are shown in Table 3.

	Estimate	St. Error	*z*-value	Pr(>|z|)
(Intercept)	−0.13	0.85	−0.16	0.88

Table 3: Fixed effects of the model of schwa in clitics.
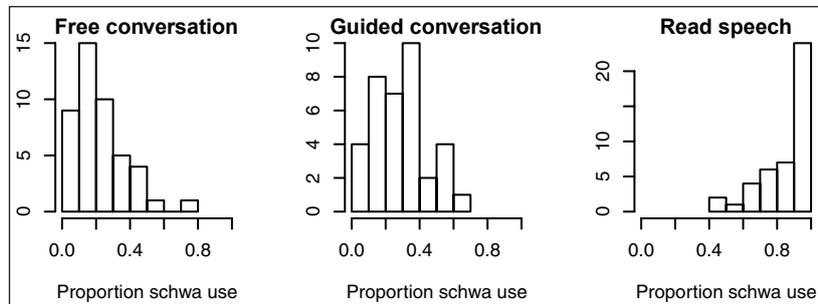
Figure 4: Proportion schwa use in clitics by speaker.

5 Discussion

Two questions motivated our study: (i) Does individual-level variation in schwa's realization exist (i.e. is schwa's optionality a speaker-level phenomenon)? (ii) Is there non-uniformity in the rate of schwa's realization across speakers (i.e. do different speakers realize schwa at different rates)? With respect to the first question, all three environments we examined reveal individual-level variation: for each context, there are at least some speakers who do not uniformly produce/omit schwa.

However, each environment also showed pockets of no variation. The Read Speech discourse type seems to encourage schwa's realization: in this discourse type, some speakers always produced schwa in clitics. To this we can add the near-universal realization of schwa in the CC___C tokens we excluded from analysis because they did not meet Côté's criteria for permitting optional schwa realization. So alongside the robust finding that individual-level variation exists, we can also see that certain phonotactic or discourse-level conditions (and perhaps other factors not examined here) can discourage or perhaps even rule out variation.

Nonetheless, our results support theories of variation that take a single speaker's grammar as the locus of the optional process. All the theories discussed in section 3 are compatible with individual-level variation, but some would have faced a severe challenge if individual-level variation had not been found. In short, this part of our study confirms that what appears to be the standard assumption in formal work on variation is on target: individual speakers (can) show variation, and theories of optionality must make multiple outputs available within a single speaker's grammar. This is not a surprising result; a vast quantity of sociolinguistic research, starting with Labov (1969), documents the pervasiveness of

variation at every level. Nevertheless, it is reassuring to know that the development of this area of phonological theory is on solid ground.

As for our second question, we found that in at least some contexts, speakers vary in the frequency with which they omit schwa – that is, the population is non-uniform in terms of the frequency of variants. This result has two immediate consequences. First, many formal analyses of optionality that model variants' frequencies (such as Hayes & Londe 2006; Anttila 2007; Kaplan 2011) model frequencies that are derived from the combined productions (or judgments) of a number of speakers. That is, they account for variants' frequencies as averaged across some population, not the frequencies produced by any one speaker. But our results suggest that this is a dangerous enterprise: if different speakers produce variants at different rates, devising a formal account that produces the average frequencies across all speakers risks producing frequencies that reflect no actual speaker's behavior. This is not to say that such analyses are never appropriate, but we should acknowledge what they do (and do not) represent and interpret them accordingly.

Furthermore, theories of optionality that aim to account for variants' frequencies of attestation must accommodate the non-uniform population revealed by our study: these theories must be able to provide different output frequencies for different speakers. Four of the seven theories examined above unquestionably meet this criterion: Stochastic OT, the Rank-Ordered Model of Eval, Markedness Suppression, and Maximum Entropy. Our results suggest that these four theories are best equipped to model phonological optionality as it manifests in French schwa: individual speakers show variation in their productions, and that variation is not constant across speakers. Returning to the dichotomy that was introduced in section 3, it seems that deriving variation through the probabilistic selection of outputs, rather than the probabilistic selection of rankings, provides the most accurate model of speaker- and population-level variation, at least for French schwa. The lone exception is S-OT, which produces variation through the probabilistic selection of rankings; we have more to say about this below.

The remaining theories – Partial Orders, Serial Variation, and Local Constraint Evaluation – are compatible with our results only under particular interpretations. If we require all speakers of a dialect to possess the same constraint ranking (this is a strong position to take, but not necessarily an unreasonable one), these theories predict uniform frequencies across speakers. Interestingly, these theories rely on the same mechanism to produce variation: some part of the constraint ranking is undetermined, and that indeterminacy can be resolved differently in different evaluations (or at different stages of a derivation in the case of SV). Without further elaboration, the frequencies these theories predict for each variant are determined solely by the likelihood of selecting a total ranking that favors that variant. If these theories are to model non-uniform populations, they require some additional machinery to provide that capability.

One simple solution, which we discussed briefly in section 3.1, is to allow speakers to vary in terms of which members of CON belong to the indeterminant part of the constraint ranking. For example, if for Speaker A the set of constraints whose ranking is undetermined contains an extra constraint against schwa in clitics compared to the grammar of Speaker B, Speaker A will produce schwa in clitics less often than Speaker B. This proposal amounts to the claim that different speakers have different grammars, which is clearly an uncontroversial assertion. We are aware of no research that has developed a system of this sort, and such an enterprise faces certain challenges. Most immediately, it relies on the presence of well-motivated constraints that can that can be included or excluded from the variable part of the ranking in a way that affects variants' frequencies without altering the set of possible outputs itself. Since the set of possible outputs and outputs' frequencies

have the same source in these theories (namely variation in the constraint ranking), they are inextricably linked, and it may not be possible to adequately disentangle them.

Another risk of this approach is that its success can remove one of the supposed advantages of this kind of theory. Frequency predictions under PO, in particular, are claimed to fall out of the analysis of optionality as byproducts; unlike theories such as MS and S-OT, PO includes no direct way of manipulating frequencies. The fact that PO happens to produce accurate output frequencies is a mark in its favor. But the suggestion in the previous paragraph robs PO of this property by arguing for manipulation of the indeterminate ranking with the explicit goal of adjusting frequency predictions. Our results suggest that such direct control over output frequencies is necessary; nonetheless, acknowledging this eliminates an argument in favor PO (and similar theories).

A second approach to non-uniform populations in PO, SV, and LCE might involve privileging some resolutions of the indeterminate ranking over others. Returning to the example above, Speaker A will produce fewer schwas in clitics than Speaker B if Speaker A's grammar is biased toward the total rankings that exclude schwa from clitics (or if Speaker B's grammar favors the total rankings that yield schwa in clitics). It is not clear to us how such a system could be implemented, but the broad strokes of some possibilities are immediately apparent. (Nagy & Reynolds 1997, working in a similar system to PO, suggest that social factors make some of the available rankings more likely than others, but they do not present a formal implementation of the idea.) First, these theories could borrow a page from the Multiple Grammars theory of variation (see Anttila 2007 and references therein), a framework that, like PO, SV, and LCE, provides multiple constraint rankings, but does so by directly stipulating what the available rankings are. That is, instead of imposing no ranking on constraints C_i and C_j, Multiple Grammars presents $C_i \gg C_j$ and $C_j \gg C_i$ as a list of possible rankings that can be chosen on any evaluation. With just this two-member set of rankings, each ranking has a 50% chance of being selected, and therefore the output each ranking produces has a predicted frequency of 50%. But by duplicating one of these rankings to produce a three-member set, as in (23), we can manipulate these frequency predictions. Under (23), the output produced by $C_i \gg C_j$ has a predicted frequency of 2/3 (because that ranking represents two of the three possibilities).

(23)　　*Rankings Available to the Grammar*
　　　　$C_i \gg C_j$
　　　　$C_i \gg C_j$
　　　　$C_j \gg C_i$

How this arrangement could be integrated into an indeterminate-ranking framework is an issue that we are not at present able to adequately address, but it is worth noting that this is essentially what sets S-OT apart from PO, SV, and LCE. All four theories produce variation through the probabilistic selection of rankings, but by doing so through continuous numerical rankings that are affected by the addition of noise, S-OT provides a more nuanced approach to output frequencies than the other theories. That is, by changing constraints' positions on the number line, S-OT can move constraints closer together or farther apart, making the likelihood of a ranking reversal once noise is added greater or smaller (respectively) without altering the inventory of available outputs. S-OT's success with non-uniform populations reinforces our argument that probabilistic selection of rankings is an inadequate approach to optionality on its own, in contrast with probabilistic selection of outputs. S-OT supplements ranking selection with another mechanism, and it is this additional piece that give it the necessary flexibility.

Alternatively, by augmenting an indeterminate-ranking framework with a system that allows constraints to favor one end of the ranking or another, we can bias the theory in favor of certain resolutions of the indeterminate ranking. For example, if C_i "likes" to be higher in the ranking than C_j , the indeterminate ranking involving C_i and C_j will be resolved into $C_i \gg C_j$ more often than the reverse. Again, we see no obvious way of incorporating this into any of the relevant theories (though Coetzee's 2002 model of crosslinguistic tendencies is very similar to this proposal), but in any case, some means of providing variation in output frequencies is needed. Adopting some continuous scale (whether through S-OT's theory of rankings or in MS's probability of violation-mark removal, e.g.) seems to be sufficient, but as the forgoing discussion indicates, it may not be the only possible approach.

A related challenge that PO-like theories face is the prediction that output frequencies are clustered around particular values – multiples of $\frac{1}{c!}$, where c is the number of variably-ranked constraints (see section 3.1). In the foregoing study we were unable to assess whether clustering is present with respect to schwa, but probing this issue is an important test of this kind of theory's viability.

In sum, theories of optionality in which variation results from the probabilistic selection of a surface form from a set of outputs, rather than probabilistic selection of constraint rankings, best reflect the patterns of schwa omission in our data. Theories that rely on the availability of multiple constraint rankings require some additional construct to accurately model output frequencies.

We wish to be clear about the limitations of our analysis, in which only a handful of factors relevant to schwa's behavior were investigated. Several salient influences could not be included in our analysis. For example, rate of speech has been shown to affect schwa (Grammont 1914; Malécot 1976; Hansen 1994; Bürki et al. 2011), as do other phonological considerations such as the number of syllables in a word (Pustka 2007). Speakers' age influences their rate of schwa omission (Malécot 1976; Hansen 1994). Lexical frequency has also been claimed to affect schwa's realization (Hansen 1994; Racine & Grosjean 2002), though Bürki et al. (2011) argue that this is an artefact of other variables. Consequently, although we aimed to weed out confounds that may give the illusion of variation while actually revealing invariant realizations in different contexts, there are some potential confounds we could not address. Similarly, we have no way of knowing whether the tokens from a particular discourse type are indeed comparable on this measure. (However, the histograms in Figures 1, 2, 3, and 4 show impressionistically that more formal discourse types encouraged more realization of [ə], a result consistent with previous literature, such as Lucci 1983; Hansen 1994; 2012; Eychenne 2006; 2009a; Pustka 2009.)

6 Conclusion

Our goal in this study was to probe how well formal theories of phonological optionality reflect the nature of the phenomena they are built to model. We found that, with respect to French schwa at least, individual-level variation does in fact exist; this is a crucial finding because most of the relevant formalisms assume that this is the case. We also found non-uniform populations in the rate of schwa production. This result has important implications for theories of variation in two significant ways. The first implication is a cautionary note: a theory that reproduces output frequencies as deduced from a corpus that includes productions from multiple speakers – or even productions from a single speaker across multiple discourse types – risks masking this non-uniformity and producing a model that matches no actual speaker's grammar. In other words, care is warranted when it comes to

this sort of analysis. Second, if a theory's goal is to capture variants' frequencies, it must be flexible enough to accommodate this non-uniformity.

Of the theories we surveyed here, S-OT, ROE, MS, and MaxEnt best reflect the empirical state of affairs uncovered by our study. All four theories are compatible with individual-level variation (in fact they all require it) and permit non-uniform populations of the sort reported here. On the other hand, PO, SV, and LCE can all produce individual-level variation but do not easily accommodate non-uniform populations. One consequence of our study, then, is that it reminds us that questions like the ones investigated above can help distinguish one theory from another; we need not rely solely on criteria such as parsimony of analysis, complexity of novel formalisms, and the ability to produce all attested variants, as important as those criteria are.

Finally, it is worth remembering that our study examined only a small segment of a single optional process. Conducting similar studies on other optional processes may very well lead to different conclusions. For example, if English speakers were to display uniformity in the rate at which flapping occurs, the empirical landscape would likely favor PO, SV, and LCE. French schwa may not be representative of all optional phenomena, and we may ultimately discover a richer typology of optionality that calls for multiple lines of analysis: competing theories of variation may not be in competition after all, with each well-suited for one or more sub-types of optionality. But these interesting empirical differences between optional processes cannot be uncovered without more studies like the one we have presented here.

Acknowledgements

Competing Interests

The authors declare that they have no competing interests.

References

Anderson, Stephen R. 1982. The analysis of French shwa: Or, how to get something for nothing. *Language* 58(2). 534–573. DOI: http://dx.doi.org/10.2307/413848

Anttila, Arto. 1997. Deriving variation from grammar. In Frans Hinskens, Roeland van Hout & W. Leo Wetzels (eds.), *Variation, change and phonological theory* (Current Issues in Linguistics 146), 35–68. Philadelphia: John Benjamins. DOI: http://dx.doi.org/10.1075/cilt.146.04ant

Anttila, Arto. 2006. Variation and opacity. *NLLT* 24(4). 893–944.

Anttila, Arto. 2007. Variation and optionality. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 519–536. Cambridge: Cambridge University Press.

Boersma, Paul. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32. 45–86.

Burki, Audrey, Mirjam Ernestus, Cédric Gendrot, Cécile Fougeron & Ulrich Hans Frauenfelder. 2011. What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech. *Journal of the Acoustical Society of America* 130(6). 3980–3991. DOI: http://dx.doi.org/10.1121/1.3658386

Coetzee, Andries. 2002. Between-language frequency effects in phonological theory. Ms. University of Massachussets, Amherst.

Coetzee, Andries W. 2004. *What it means to be a loser: Non-optimal candidates in Optimality Theory*. Amherst: University of Massachussets dissertation.

Coetzee, Andries W. 2006. Variation as accessing 'non-optimal' candidates. *Phonology* 23(3). 337–385. DOI: http://dx.doi.org/10.1017/S0952675706000984

Côté, Marie-Hélène. 2001. *Consonant cluster phonotactics: a perceptual approach*. Cambridge, MA: MIT dissertation. DOI: http://dx.doi.org/10.1017/S0959269507002827

Côté, Marie-Hélène. 2012. Laurentian French (Quebec): Extra vowels, missing schwas and surprising liaison consonants. In Randall Gess, Chantal Lyche & Trudel Meisenburg (eds.), *Phonological variation in French: Illustrations from three continents* (Studies in Language Variation), 236–274. John Benjamins.

Côté, Marie-Hélène & Geoffrey Stewart Morrison. 2007. The nature of the schwa/zero alternation in French clitics: Experimental and non-experimental evidence. *Journal of French Language Studies* 17. 159–186.

Dell, François. 1973. *Les règles et les sons: Introduction à la phonology générative*. Paris: Hermann. Translated in 1980 by Catherine Cullen as *Generative Phonology and French Phonology*. New York: Cambridge University Press.

Dell, François. 1977. Paramètres syntaxiques et phonologiques qui favorisent l'épenthèse de schwa en français moderne. In Christian Rohrer (ed.), *Actes du colloque franco-Allemand de linguistique théorique,* vol. 39, 141–153. Tubingen: Max Niemeyer Verlag.

Dell, François. 1980. *Generative phonology and French phonology*. Cambridge: Cambridge University Press. Translated by Catherine Cullen.

Durand, Jacques, Bernard Laks & Chantal Lyche. 2002. La phonologie du français contemporain: Usages, variétés et structure. In Claus D. Pusch & Wolfgang Raible (eds.), *Romanistische korpuslinguistik: Korpora und gesprochene sprache,* 93–106. Tubingen: Gunter Narr Verlag.

Durand, Jacques, Bernard Laks & Chantal Lyche. 2009a. Le projet PFC: Une source de données primaires structurées. In Jacques Durand, Bernard Laks & Chantal Lyche (eds.), *Phonologie, variation et accents du français*, 19–61. Paris: Hermès.

Durand, Jacques, Bernard Laks & Chantal Lyche (eds.). 2009b. *Phonologie, variation et accents du français*. Paris: Hermès.

Durand, Jacques, Catherine Slater & Hilary Wise. 1987. Observations on schwa in southern French. *Linguistics* 25(2). 983–1004. DOI: http://dx.doi.org/10.1515/ling.1987.25.5.983

Encrevé, Pierre. 1988. *La liaison avec et sans enchaînement: Phonologie tridimensionelle et usages du français*. Paris: Editions du Seuil.

Eychenne, Julien. 2006. *Aspects de la phonologie du schwa dans le français contemporain: Optimalité, visibilité prosodique, gradience*: Université de Toulouse-Le Mirail dissertation.

Eychenne, Julien. 2009a. La pronunciation du français au Pays Basque. In Jacques Durand, Bernard Laks & Chantal Lyche (eds.), *Phonologie, variation et accents du français*, 231–258. Paris: Hermès.

Eychenne, Julien. 2009b. Une variété français conservatrice en Languedoc. In Jacques Durand, Bernard Laks & Chantal Lyche (eds.), *Phonologie, variation et accents du français*, 259–284. Paris: Hermès.

Fougeron, Cécile, Cedric Gendrot & Audrey Burki. 2007. On the acoustic characteristics of French schwa. In *Proceedings of ICPhS XVI*, 941–944.

Gess, Randall, Chantal Lyche & Trudel Meisenburg (eds.). 2012. *Phonological variation in French: Illustrations from three continents* (Studies in Language Variation). Amsterdam: John Benjamins. DOI: http://dx.doi.org/10.1075/silv.11

Goldwater, Sharon & Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the workshop on variation within Optimality Theory*, 113–122. Stockholm University.

Grammont, Maurice. 1914. *Traité pratique de prononciation française*. Paris: Delagrave.

Hansen, Anita Berit. 1994. Etude du E caduc – stabilisation en cours et variations lexicales. *Journal of French Language Studies* 4. 25–54. DOI: http://dx.doi.org/10.1017/S0959269500001964

Hansen, Anita Berit. 1997. Le nouveau [ə] prépausal dans le français parlé à Paris. In Jean Perrot (ed.), *Polyphonie pour Iván Fónagy: Mélanges offert en homage à Iván Fónagy*, 173–198. Paris: Harmattan.

Hansen, Anita Berit. 2012. A study of young Parisian speech: Some trends in pronunciation. In Randall Gess, Chantal Lyche & Trudel Meisenburg (eds.), *Phonological variation in French: Illustrations from three continents* (Studies in Language Variation), 152–172. John Benjamins.

Hayes, Bruce & Zsuzsa Cziráky Londe. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23. 59–104.

Howard, Irwin. 1973. *A directional theory of rule application in phonology*. Indiana University Linguistics Club.

Jäger, Gerhard. 2007. Maximum Entropy models and Stochastic Optimality Theory. In Annie Zaenen, Jane Simpson, Tracy Holloway King, Joan Maling Jane Grimshaw & Chris Manning (eds.), *Architectures, rules, and preferences: Variations on themes by Joan W. Bresnan*, 467–479. Stanford: CSLI. DOI: http://dx.doi.org/10.1515/LING.2006.031

Jäger, Gerhard & Anette Rosenbach. 2006. The winner takes it all – almost: Cumulativity in grammatical variation. *Linguistics* 44(5). 937–971.

Jesney, Karen. 2007. The locus of variation in weighted constraint grammars. Poster presented at the Workshop on variation, gradience and frequency in phonology. Stanford, CA: Stanford University. July 2007.

Kaplan, Aaron 2011. Variation through markedness suppression. *Phonology* 28(3). 331–370. DOI: http://dx.doi.org/10.1017/S0952675711000200

Kaplan, Aaron. to appear. Local optionality with partial orders. *Phonology*.

Kimper, Wendell. 2011. Locality and globality in phonological variation. *Natural Language and Linguistics Theory*. 423–465.

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45(4). 715–762.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Harmonic Grammar – a formal multi-level connectionist theory of linguistic well-formedness: An application. In *Proceedings of the twelfth annual conference of the cognitive science society*, 884–891. Cambridge, MA: Lawrence Erlbaum.

Léon, Pierre R. 1987. E caduc: Facteurs distributionnels et prosodiques dans deux types de discours. In *Proceedings of the eleventh international congress of phonetic sciences*, vol. 3, 109–112.

Lucci, Vincent. 1983. *Etude phonétique du français contemporain à travers la variation situationnelle: Débit, rythme, accent, intonation, ə muet, liaisons, phonèmes*. Grenoble: Université des langues et lettres de Grenoble.

Malécot, André. 1976. The effect of linguistic and paralinguistic variables on the elision of the French mute-e. *Phonetica* 33. 93–112. DOI: http://dx.doi.org/10.1159/000259716

Munro, Pamela & Jason Riggle. 2004. Productivity and lexicalization in Pima compounds. In Marc Ettlinger, Nicholas Fleisher & Mischa Park-Doob (eds.), *Proceedings of the 30th annual meeting of the Berkeley Linguistics Society: Special session on the morphology of Native American languages*, 114–126. Berkeley: Berkeley Linguistics Society.

Nagy, Naomi & Bill Reynolds. 1997. Optimality Theory and variable word-final deletion in Faetar. *Language Variation and Change* 9. 37–55. DOI: http://dx.doi.org/10.1017/S0954394500001782

Nevins, Andrew & Bert Vaux. 2008. Introduction: The division of labor between rules, representations, and constraints in phonological theory. In Bert Vaux & Andrew Nevins (eds.), *Rules, constraints, and phonological phenomena*, 1–19. Oxford: Oxford University Press.

Poiré, François. 2009. Le français canadien en milieu minoritaire: Le cas du Sud-Ouest ontarien. In Jacques Durand, Bernard Laks & Chantal Lyche (eds.), *Phonologie, variation et accents du français*, 153–173. Paris: Hermès.

Prince, Alan & Paul Smolensky. 1993/2004. Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick and University of Colorado, Boulder. Published 2004, Malden, MA: Blackwell.

Pustka, Elissa. 2007. *Phonologie et variétés en contact: Aveyronnais et Guadeloupéens à Paris*. Tubingen: Narr.

Pustka, Elissa. 2009. PFC et la phonologie du français en Vendee. In Jacques Durand, Bernard Laks & Chantal Lyche (eds.), *Phonologie, variation et accents du français*, 307–335. Paris: Hermès.

Racine, Isabelle & François Grosjean. 2002. La production du e caduc facultatif est-elle prévisible? un début de réponse. *Journal of French Language Studies* 12(3). 307–326. DOI: http://dx.doi.org/10.1017/S0959269502000340

Reynolds, William Thomas. 1994. *Variation and phonological theory*. Philadelphia: University of Pennsylvania dissertation.

Riggle, Jason & Colin Wilson. 2005. Local optionality. In *NELS 35*, 539–550.

Selkirk, Elisabeth. 1978. The French foot: On the status of "mute" e. *Studies in French Linguistics* 1(2). 141–150.

Smolensky, Paul & Géraldine Legendre. 2006. *The harmonic mind*. Cambridge, MA: The MIT Press.

Steriade, Donca. 1999a. Alternatives to the syllabic interpretation of consonantal phonotactics. In Osamu Fujimura, Brian D. Joseph & Bohumil Palek (eds.), *Proceedings of the 1998 linguistics and phonetics conference*, 205–242. The Karolinum Press.

Steriade, Donca. 1999b. Phonetics in phonology: The case of laryngeal neutralization. In Matthew K. Gordon (ed.), *UCLA Working Papers in Linguistics*, vol. 2, 25–146.

Steriade, Donca. 2001. Directional asymmetries in place assimilation. In Elizabeth Hume & Keith Johnson (eds.), *The role of speech perception in phonology*, 220–250. San Diego: Academic Press.

Tranel, Bernard. 1987. *The sounds of French*. Cambridge: Cambridge University Press. DOI: http://dx.doi.org/10.1017/CBO9780511620645

Van Eibergen, Joëlle. 1991. Evaluation quantative de l'alternance phonétique du /schwa/: Importance de l'entourage consonantique. In *Actes du XIIème congrès international des sciences phonétiques*, 150–152. Aix-en-Provence: Université de Provence.

Van Eibergen, Joëlle. 1992. Effacements et maintiens du /schwa/ bifide. In *Bulletin de la communication parlée no. 2*, 115–124. Institut de la Communication Parlée.

Van Eibergen, Joëlle & Rabia Belrhali. 1994. Étude statistique du /schwa/. In Chantal Lyche (ed.), *French generative phonology: Retrospective and perspectives*, 277–287. Salford: Association for French Language Studies.

Vaux, Bert. 2008. Why the phonological component must be serial and rule-based. In Bert Vaux & Andrew Nevins (eds.), *Rules, constraints, and phonological phenomena*, 20–60. Oxford: Oxford University Press. DOI: http://dx.doi.org/10.1093/acprof:oso/9780199226511.003.0002

Walter, Henriette. 1982. *Enquête phonologique et variétés régionales du français*. Paris: Presses Universitaires de France.

]u[*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.   OPEN ACCESS