Direct Mapping of Acoustics to Phonology: On the lexical encoding of front rounded vowels in L1 English-L2 French acquisition

Isabelle Darcy, Laurent Dekydtspotter, Rex A. Sprouse,
Justin Glover, Christiane Kaden, Michael McGuire, and John H.G. Scott
(Indiana University)

Corresponding author:
Isabelle Darcy
Dept. of Second Language Studies
Indiana University
Memorial Hall 315, 1021 E Third St,
Bloomington, IN 47405
USA
idarcy@indiana.edu

Author's copy (prefinal)
Accepted for publication in Second Language Research, August 2011.

Direct Mapping of Acoustics to Phonology:
On the lexical encoding of front rounded vowels in
L1 English-L2 French acquisition

Abstract

It is well known that adult US-English-speaking learners of French have difficulties acquiring high /y/-/u/ and mid /œ/-/ɔ/ front versus back rounded vowel contrasts in French. This study examines the acquisition of these French vowel contrasts at two levels: phonetic categorization and lexical representations. An ABX categorization task revealed that both advanced and intermediate learners failed to categorize /œ/ versus /ɔ/ and /y/ versus /u/ as native speakers of French do, although performance on the /y/-/u/ contrast was better than on the /œ/-/ɔ/ contrast in all contexts. On a lexical decision task with repetition priming, advanced learners and native speakers produced no (spurious) response time facilitations for /y/-/u/ and /œ/-/ɔ/ minimal pairs; however, in intermediate learners, the decision for a word containing /y/ was speeded by hearing an otherwise identical word containing /u/ (and vice-versa), suggesting that /u/ and /y/ are not distinguished in lexical representations. Thus, while it appears that advanced learners encoded the /y/-/u/ and /œ/-/ɔ/ contrasts in the phonological representations of lexical items, they gained no significant benefit on the categorization task**.** This dissociation between phonological representations and phonetic categorization challenges common assumptions about their relationship and supports a novel approach we label *Direct Mapping from Acoustics to Phonology* (DMAP).

I. Introduction

Research on second language (L2) sound systems has focused on two distinct levels of investigation: acquisition of non-native phone categories and acquisition of abstract phonological representations. A number of models, including the Speech Learning Model (Flege, 1995), the Perceptual Assimilation Model (Best, 1995; PAM-L2, Best and Tyler, 2007), and the Native Language Magnet Model (Kuhl and Iverson, 1995), have focused on the problem of categorization. According to these models, under specific conditions, the configuration of the native language (L1) phonetic space induces the classification of target (second) language phones as (potentially target-deviant) instances of a category in the learner's L1 (e.g. through "single-category assimilation" as described by Best, 1995; Best, McRoberts and Goodell, 2001; Best, McRoberts and Sithole, 1988). In such cases, acquiring the relevant target-language categories can pose a severe challenge. Another avenue of research focuses on the development of interlanguage phonological representations that determine phonotactics, stress, tonal or intonation patterns, segmental inventories, syllable structure constraints, etc. (Altenberg and Vago, 1983; Archibald, 1993; 1998; Broselow, 1987; Broselow, Chen and Wang, 1998; Eckman, 1977; 1987; Young-Scholten, 2004). Some L2 phonological researchers also claim that perception is constrained by the features used to specify the L1 phonological segment inventory. For Brown (1998, 2000), only the features strictly necessary to specify the L1 phonological segment inventory are available. Hancin-Bhatt (1994) argues that accurate perception of phone contrasts correlates with feature prominence: Features specified in more segments of the L1 inventory allow better discrimination. Beyond the segmental level, work on the Prosodic Transfer Hypothesis argues that the L1 prosodic grammar filters the integration of segments, explaining aspects of L2 phonological productions under various strategies (Goad and White, 2006; 2008; Goad, White and Steele, 2003).

Research on the L2 acquisition of new phonetic contrasts widely (albeit, implicitly) assumes that the development of a new category in the perceptual space constitutes the first stage of the acquisition of (one or more) new phonemes, in analogy with L1 acquisition (Maye, 2000; Maye, Werker and Gerken, 2002). This mirrors commonly accepted views of word recognition, where the output of phonetic categorization (in which irrelevant variation has been discarded) is the input to phonology. As an example of this assumption, Pallier, Bosch and Sebastián-Gallés (1997), observing unreliable discrimination of Catalan words containing [ɛ] from minimally different words containing [e] in Spanish-dominant bilinguals, argue that these bilinguals did not generally establish new categories for Catalan /ɛ/, despite exposure from an early age. Studying a very similar population, Pallier, Colomé and Sebastián-Gallés (2001) observed spurious repetition priming for Catalan /e/-/ɛ/ minimal pairs, which they interpreted as evidence that Spanish-dominant bilinguals treated such word pairs as homophones, unlike Catalan native speakers. Even though the lexical homophony interpretation is not the only one possible (the same results could be due to the listeners' inability to auditorily *distinguish* the minimal pairs in the first place), these authors clearly relate the lacking lexical distinction to the bilinguals' failure to establish distinct categories for the Catalan phonemes. However, other researchers have also found cases where L2 learners seem to have knowledge of lexical contrasts despite unreliable performance on discrimination tasks (Cutler, Weber and Otake, 2006; Escudero, Hayes-Harb and Mitterer, 2008; Hayes-Harb and Masuda, 2008; Weber and Cutler 2004). The same assumption has led these researchers to hypothesize *ersatz* lexical representations and mechanisms of lexical encoding that rely on other (i.e., metalinguistic) sources of knowledge.

This paper challenges the assumption that phonetic distinctions must precede phonological contrasts in lexical representations by examining the L2 acquisition of two vowel contrasts that occur in French, but not in English. In French, the high front rounded vowel /y/ contrasts with the high back rounded vowel /u/ and the mid front rounded vowel /œ/ contrasts with the mid back rounded vowel /ɔ/, whereas in English neither front rounded vowel occurs. We examine in tandem the degree to which intermediate and advanced English-speaking learners can categorize the front versus back, rounded vowel contrasts /y/-/u/ and /œ/-/ɔ/ in non-words, as well as these same learners' lexical representations of /y/-/u/ and /œ/-/ɔ/ minimal pairs. Thus, we consider these contrasts at two levels: segmental phonetic categorization and phonemic representations in the lexicon. Segmental categorization is examined with an ABX task, and lexical representations with a lexical decision task in the repetition priming paradigm.

A research program argues that phonological contrasts, processes and alternations require discrete features organized according to a hierarchy (Clements, 1985, 2001, 2003, 2009; Clements and Hume, 1995, Keyser and Stevens, 1994; McCarthy, 1988; Halle, 1992, Dresher, 2009, 2010; and many others). Features and their geometry are transduced from the sensory-motor system into the phonological domain (Keyser and Stevens, 1994). Features "correspond to articulatory regions with relatively stable acoustic properties" (Clements, 2009:19). These discrete phonological features receive context-dependent category definitions on a continuum (Kuhl and Iverson, 1995). The hierarchy minimizes redundancy, expresses universal tendencies and underlies phonological processes. A feature must be relevant either to lexical contrasts, to phonological patterns or alternations, or to phonetic realizations, in order to be selected in the acquisition of a language (Clements, 2001). The hierarchy of features, supplemented with phonetic cues to phonological parameters, allows phonological acquisition to abstract away from irrelevant phonetic details (Dresher 1999; Dresher and Kaye 1990). In contrast, phonetically grounded approaches forego innate features grounded in the sensory motor system (Ohala, 1995; Steriade 1999; *inter alia*). Mielke (2008) views features as emergent properties of the signal. Clements (2001: 84-85) notes:

> "While it is possible that the hierarchy is simply given as such in universal grammar, it is not unreasonable to suppose that it can be recovered, at least in large part, from the speaker's linguistic experience through massive exposure to data allowing a calculation of relative phoneme frequencies and other phenomena related to feature accessibility."

For (adult) L2 acquisition, a hierarchy of features selected or abstracted in L1 acquisition predates exposure to L2 input. Indeed, this selection mediates L2 acquisition (Brown, 1998, 2000; Hancin-Bhatt, 1994), as does prosodic selection (Goad and White, 2006; 2008; Goad, White and Steele, 2003).

In French, front rounded vowels receive [front] + [round] specifications, whereas back rounded vowels receive [back] + [round] specifications. Following Clements and Hume (1995), we will assume that the features [round], [front] and [back] are reducible to [labial], [coronal] and [dorsal] V-place specifications of articulators, enforcing constrictions due to lip, tongue blade, and tongue body respectively. The contrast /y/-/œ/ is mediated by tongue height under an aperture/vowel height node. Hence, as the literature (e.g. Levy, 2009a; Flege, 1987; Gottfried, 1984) suggests, we expect neither /y/-/u/ nor /œ/-/ɔ/ phonological contrasts in the initial state of L2 phonological acquisition of French by native speakers of English. (We discuss this in detail in Section II.) Following Schwartz and Sprouse's (1994; 1996) Full Transfer/Full Access model motivated by morphosyntax and adopted for phonology by Archibald (1998), we assume that the L1 phonological system (or phonological grammar) constitutes the initial phonological state in L2 acquisition. English has no contrast between front and back rounded vowels: rounding merely enhances the front versus back contrast (Clements, 2001). Acoustically, however, in many varieties of US English, /u/ comes close to a front vowel, especially in coronal contexts (Levy, 2009a, b; Hillenbrand et al., 2001, Strange et al., 2007). The French values for the /y/-/u/ contrast therefore are found at the margins of the category definition for English /u/. Hence, due to the transfer of the feature specifications that carve out the English inventory of segments, rounded vowels (front as well as back) in the input will be assigned the [dorsal] specification in lexical representations in the initial stages of English-based interlanguage development. Borrowing from historical linguistics, we refer to this L1-based reinterpretation of the phonological content of the target-language input as a "merger" in order to distinguish this process from perceptual assimilation at the level of phonetic categories.

Following the logic of Pallier et al. (2001), phonological merger routinely leads to "spurious homophony" in the interlanguage lexicon. This leads us to ask whether phonological merger can be overcome and if so, how lexical representations can be revised to reflect the new phonological state. Perceptual assimilation (Best, 1995) in which foreign speech sounds are treated as exemplars of L1 phonetic categories also characterizes aspects of the initial state of L2 sound systems. This also invites the question of the degree to which L2 learners can recover from single-category assimilation, crucially modulated by consonantal context, in the process of L2 acquisition (Levy and Strange, 2008; Levy, 2009 a, b). Levy and Strange show that experienced learners exhibited only a marginal improvement over inexperienced learners: Experienced learners' categorization of /y/-/u/ was not affected by the consonantal context, but rates of categorization errors remained similar to inexperienced learners. If learners' lexical representations recover from merger, yet learners' categorization of the relevant phones exhibits little benefit, this would challenge the assumption that the establishment of new categories (i.e. recovery from single-category assimilation) constitutes the first step in the L2 acquisition of a phonological contrast.

In Section II we review the literature on category formation in L2 sound perception and on the lexical representation of L2 phonological contrasts. In Section III, we examine the acquisition problem posed by French rounded vowels for native speakers of English. We consider what is strictly necessary for phonological development to occur without recourse to other information sources and propose *Direct Mapping from Acoustics to Phonology* (*DMAP*) as a possible mechanism underlying phonological development. Empirical evidence for DMAP is presented in Sections IV and V, where we present evidence of phonological merger (and recovery from merger) in the lexicon and examine learners' categorization ability in relation to lexical knowledge involving the same contrast. A general discussion of some of the implications of DMAP follows in Sections VI and VII.

II. Category formation and lexical encoding of contrasts

2.1 Categorization and phonetic decoding

In spoken language perception research, a large number of studies have documented categorical perception (i.e. indicating the presence of categories as well as their boundaries), whereby categorization performance (identification of categories on a speech-sound continuum) predicts performance on discrimination tasks (e.g. Fujisaki and Kawashima, 1971; Liberman, Harris, Hoffman and Griffith, 1957; Pisoni, 1973). For speech, it is generally assumed that categorization of acoustic stimuli represents a basic and automatic step in speech processing. Categorical perception of phonemes results in minimizing perceived differences between sounds along one or more dimensions within the category boundaries and enhancing perceived differences on those dimensions across the boundaries. An acoustic change is perceived most clearly when it crosses a phoneme boundary (Dehaene-Lambertz, 1997). This stretching and squeezing of the perceived distance between stimuli reflects the influence that categories have on the perceptual similarity of acoustically equidistant stimuli, as demonstrated for the speech categories /r/ and /l/ by Iverson et al. (2003). This mechanism is in place very early in life (Eimas, Siqueland, Juczyk and Vigorito, 1971; Eimas, 1974). The linguistic environment rapidly modifies initial capacities and carves out language-specific perceptual boundaries (Maye, Werker and Gerken, 2002; Werker and Curtin, 2005; Kuhl, Williams, Lacerda, Stevens and Lindblom, 1992).

2.2 Modification of the categorization space

Categories remain modifiable to some extent later in life, and the ability to form new categories, at least momentarily, remains present across the life span (Maye, 2000; Maye, Werker and Gerken, 2002). In order to acquire a new category or to modify an existing category boundary, perceivers must attune to appropriate perceptual dimensions (Francis, Baldwin and Nusbaum, 2002; Francis and Nusbaum 2002), in order to match them to definitional criteria. Researchers have documented short-term shifts in consonant category boundaries through exposure to ambiguous sounds linked to lexical items (McQueen, Cutler and Norris, 2003; 2006; Eisner and McQueen, 2005; Evans and Iverson, 2002). Robust category formation seems dependent on high phonetic variability (for /r/-/l/, Lively, Logan and Pisoni, 1993; Lively, Pisoni, Yamada, Tohkura and Yamada, 1994; Logan, Lively and Pisoni, 1991; 1993). However, the long-term effects of such adjustments obtained in the laboratory are unclear for L2 acquisition. Training studies involving difficult contrasts seem much less successful at achieving long-term success (Bradlow, Pisoni, Akahane-Yamada and Tohkura, 1997).

Vowel continua are perceived in a less categorical fashion than stop consonant ones, for which discrimination peaks are usually more sharply defined. Discrimination is higher for vowels even within category boundaries (Stevens, Liberman, Studdert-Kennedy and Öhman, 1969), probably because (as pointed out by Pisoni, 1973) two types of memory information are simultaneously available for vowels (auditory and phonetic). To date, however, it remains an open question whether it is easier to *acquire* target-like vowel categories than target-like consonant categories for adult learners (Pallier, Bosch and Sebastián-Gallés, 1997). The development of new segmental categories that would interfere with the perceptual space optimized for L1 categories appears to severely tax resources, potentially inhibiting learning. Pallier et al. (1997) showed that Spanish native speakers whose exposure to Catalan began at the age of 4 years and became highly proficient in Catalan still experienced difficulty discriminating words containing [ɛ] from minimally different words containing [e] in Catalan. A synthetic vowel continuum of acoustically equidistant [ɛ] and [e] values was used to assess discrimination with an AX task and to assess identification with a classification task asking the question "Does the vowel sound more like in [pɛra] or in [pera]?". Group results showed that Spanish-dominant bilinguals had a very flat identification curve, representing random performance, along with a discrimination function that failed to show the typical peak around the category boundary. By contrast, Catalan-dominant bilinguals exhibited an identification curve indicating a clear labeling of two categories at the extreme ends of the continuum, and excellent discrimination around the category boundary (Pallier, Bosch and Sebastián-Gallés, 1997). These results clearly indicate that many proficient L2-learners of Catalan, despite their early exposure, do not successfully establish segmental categories for the Catalan /ɛ/ or /e/ and use their native Spanish /e/ to categorize both members of the pair.

Turning to rounded vowels, previous studies have shown that (US-) English-speaking acquirers of French have more difficulty discriminating /y/-/u/ pairs than /i/-/y/ pairs (e.g., Flege, 1987; Gottfried, 1984; see also Flege and Hillenbrand, 1984, on production). This was independent of context. However, Polka (1995) observed that (US-)English-speaking listeners without any experience with German exhibit native-like discrimination of the /y/-/u/ tense vowel contrast in German, but not of the corresponding /ʏ/-/ʊ/ lax vowel contrast. Error rates were extremely low, below 10% (which she qualified as native) for the tense contrast and between 10% and 15% for the lax contrast (only 2 of the 10 listeners had error rates lower than 10%). While this performance seems surprisingly good, task characteristics might have played a role in these near-ceiling accuracy levels. For example, in the task, an effort was made to reduce acoustic variability: Stimuli were all produced by the same male voice, and there were only 6 tokens per contrast, embedded in monosyllabic /dVt/ syllables (coronal context) recorded in citation form; furthermore, time pressure and memory load were very low. Thus, both auditory and phonetic memory codes might have entered into the discrimination response (see Pisoni, 1973), thereby producing very high performance on the /y/-/u/ contrast. However, when confronted with different sources of variation, this discrimination ability is not very robust. Indeed, the stability of discrimination across contexts or despite phonetic variability can be viewed as a defining feature of the presence of a segmental category. Crucially, these data suggest that discrimination can be excellent even when no new categories have been acquired.

Levy and Strange (2008) investigated the perception of a series of French contrasts in different consonantal contexts in L1-English learners of French. In their task, multisyllabic non-words appeared in sentences. Levy and Strange varied the consonantal context systematically and used three different female voices. They observed better general performance in advanced learners of French than in inexperienced learners, but error rates on the /y/-/u/ contrast remained comparable to inexperienced learners at 25%. A development was found, however, modulated by context. Levy and Strange (2008) found that listeners with no French experience made more categorization errors for the /y/-/i/ contrast than for the /y/-/u/ contrast in labial contexts, a finding consistent with data in Levy (2009b). Indeed, the consonantal environment changes the formant patterns of vowels (Hillenbrand, Clark and Nearey, 2001).

Levy (2009a) examined the role played by allophonic variation in cross-language perceptual assimilation: Respondents whose French experience ranged from none to advanced classified French vowels in terms of six US English vowel categories and gave a goodness rating for each. In bilabial contexts, classification of /y/ was affected by experience: Assimilations to English /i/ in subjects with no experience were eliminated in favor of English /u/ with more exposure to French. The fronting of /u/ in coronal contexts in English is clearly shown to play a role in this cross-language perceptual assimilation task, interfering with categorization on an ABX task (Levy, 2009b). Experienced learners, however, had recovered from this particular error with /i/-/y/. Yet, despite having learned the contextual variation within vowel categories, they still displayed relatively high rates of categorization errors on the /y/-/u/ contrast. The error rate for /y/-/u/ in Levy and Strange's (2008) study was around 25% to 30%, constituting a much lower performance than in Polka's (1995) perception task. Connected speech places higher demands on memory, because (potential) words must be extracted from the speech stream. These stimuli were much longer, placing higher demand on the auditory memory buffer. They also contained more acoustic variability, so that matching strategies that could make use of low-level acoustic similarity were not readily available.

The /œ/-/ɔ/ contrast has received less attention in the literature. In Levy (2009a), the range of English categories selected as matching the French /œ/ was wide (/ɚ/, /ʌ/, /ʊ/, /ɔ/ and /u/), and varied by context and experience. As mentioned above, Levy (2009b) examined these learners' categorization in an AXB task in light of the perceptual assimilation data in Levy (2009a). The respondents with no to moderate exposure to French exhibited considerably more categorization errors in the /œ/-/ɔ/ contrast in coronal contexts than in bilabial contexts. This seems to correlate with the fact that /œ/ tended to assimilate to /u/ in bilabial contexts. At the same time, there were more /œ/-/y/ categorization errors in coronal contexts than in bilabial contexts. Unsurprisingly, errors rates were much reduced in advanced learners.

All in all, it appears that the front vs. back rounded vowel contrasts of French present serious difficulties for US-English native speakers, particularly in coronal context, and when the categorization task at hand is demanding. There are no data yet regarding the lexical encoding of these vowel contrasts in learners of French.

2.3 Lexical encoding of a contrast

Prima facie, it seems reasonable to assume that an (L1 or L2) learner who cannot reliably distinguish between two target-language phones will collapse them into a single category and consequently fail to lexically encode the contrast. Consequently, minimal pairs in the target language will correspond to "spurious" homophone in the learner's interlanguage lexicon. Therefore, the bulk of research on L2 acquisition of phonology has dealt with the acquisition of category distinctions. A few studies investigating consequences of perceptual mis-categorizations at the lexical level support a one-to-one mapping between reliable phonetic distinction and lexical contrast (e.g. Pallier, Bosch and Sebastián-Gallés 2001; Ota, Hartsuiker and Haywood, 2009; Pater 2003). Following up on Pallier et al. (1997), described in section 2.2, Pallier, Colomé and Sebastián-Gallés (2001) used the paradigm of lexical decision, in which the subject was asked "is the item a word?", with long-term repetition priming to investigate the lexical encoding of the /ɛ/-/e/ contrast in Catalan. Upon hearing a list of words and non-words, Catalan-dominant bilinguals performed the decision task more quickly on actual words when the given item had already been presented (/pera/ - /pera/, repetition priming) than when it was preceded by a minimally different item (/pɛra/ - /pera/). The Spanish-dominant participants listening to the same stimuli exhibited repetition priming effects equally for minimal pairs and actual repetitions. Pallier, Colomé and Sebastián-Gallés (2001) argue that both members of such a minimal pair were encoded as homophones in the Spanish-dominant participants' mental lexicon (See also Dufour, Nguyen and Frauenfelder, 2007; Sebastián-Gallés, Echeverría and Bosch, 2005). It could be argued that the effect is due to a lack of discrimination of the minimal pairs. That is to say, in these online listening tasks, the L2 respondents might not have detected the difference between /pera/ and /pɛra/, even if they had developed contrasting lexical representations. As a result, they would have accessed the same lexical representations regardless of the vowel in the auditory input (Escudero, Hayes-Harb and Mitterer, 2008). In any case, these data are compatible with the claim that the acquisition of a robust L2 phonetic category is a pre-requisite for the acquisition of target-like lexical representations.

Recently, Ota, Hartsuiker and Haywood (2009) offered an experimental methodology designed to avoid any problem caused by auditory stimuli. In their task, respondents judged the semantic relatedness of two words presented visually. For instance, Japanese-English learners (but not English native speakers) judged pairs such as LOCK/HARD and ROCK/KEY to be semantically related at much greater rate than control items. This shows that reading LOCK had activated the semantic network of ROCK and vice-versa. In this experiment, online auditory misperception cannot account for the observed cross-lexical activation of /r/-/l/ minimal pairs. As Ota, Hartsuiker and Haywood (2009: 267) state: "the lexicon of late bilinguals indeed fails in completely separating L2 lexical entries that involve nonnative phonological contrasts." The data just reviewed thus appear consistent with the claim that the ability to establish phonetic categories is a pre-requisite for encoding the contrast lexically.

Contrary to the "categories first" view, however, a body of research ("lexicon first") provides evidence that lexical contrasts can be made by proficient L2 learners even when the relevant L2 phones are not (yet) well discriminated (Cutler, Weber and Otake, 2006; Escudero, Hayes-Harb and Mitterer, 2008; Hayes-Harb and Masuda, 2008; Weber and Cutler 2004). That is to say, the learner can somehow establish a lexical contrast, although he or she cannot reliably categorize two phonemes of the target-language as different. For instance, on the basis of an eye-tracking experiment, Weber and Cutler (2004) show that lexical encoding of a difficult contrast is possible, even if the categorization of the contrast is not robust. In highly proficient Dutch-English learners, syllables with /æ/ activated minimally contrasting syllables with /ɛ/ ([pæn]-[pɛn]). However, there was no mutual activation in the other direction (/ɛ/ only activated /ɛ/ syllables), which led Weber and Cutler to conclude that these sounds do not lead to the creation of homophones at the lexical level. The authors suggested a potential effect of orthography or metalinguistic knowledge. This effect of orthography has been further supported by Escudero et al. (2008), who showed that exposing a group to the orthographic form of the words (names for non-objects) to be learned allowed for quasi-immediate recognition of the non-objects in a visual world paradigm. By contrast, the experimental group not exposed to the orthographic form of the non-object names was slower in identifying the intended referent.

Hayes-Harb and Masuda (2008) investigated English-Japanese learners' lexical representations of Japanese words containing geminate consonants, which are unattested in English. Naive English native speakers and learners with one year of exposure to Japanese, as well as native Japanese speakers, participated in a picture-matching task involving pseudo-words. During a training phase, participants were required to learn the association between 12 pictures and their (invented) brand names. Critical minimal pairs involved singleton versus geminate consonants. In a listening task where this pairing was tested (names were paired with correct and incorrect pictures), learners with one year of exposure to Japanese were excellent, and not significantly different from Japanese natives. In a naming task (in which participants were to produce the correct name for each picture), however, the learners did not make as robust a distinction as native speakers. The learners exhibited sharp boundaries, but those were not target-like. Echoing Cutler, Weber and Otake (2006), Hayes-Harb and Masuda (2008: 28) conclude that learners might lexically encode “a geminate /tt/ consonant as /t*/, where the ‘*’ might mean ‘sounds different from /t/’, even if they have not yet determined specifically how /t/ and /t*/ differ”. Hayes-Harb and Masuda (2008) state that the distinctions they find might not reflect full phonetic and phonological acquisition in some learners. Clearly, if these learners lexically represent geminates, the acoustic and articulatory targets for geminates do not appear to be target-like. However, Hayes-Harb and Masuda’s proposed explanation (as well as Weber and Cutler’s, 2004) requires specific assumptions about the nature of the lexical representations and mechanisms of lexical encoding, involving accessibility of metalinguistic information.

While this is of course a possible explanation, more research is needed to specify the extent to which metalinguistic (e.g. visual) and orthographic knowledge can support the development of an interlanguage phonological system including lexical representations. Our goal is not to deny any supportive role of orthographic or metalinguistic knowledge in second language development; rather, we want to contribute to the debate by investigating the degree to which a purely grammatical explanation is supported.

In sum, recent results challenge the prevailing view that the acquisition of distinct categories drives the establishment of the relevant lexical contrast: the segmental categorization data seem dissociated from the ability to establish lexical contrasts. The nature of these contrasting lexical representations is unclear, and the mechanisms that could lead to such a contrast are rather mysterious. In view of this, we investigate the relationship between category acquisition in the perception of rounded vowels and the acquisition of rounded vowel lexical contrasts. We investigate learners’ categorization of target-language contrasts on an ABX task as well as their lexical encoding of those contrasts on a lexical decision task that parallels the one used by Pallier, Colomé and Sebastián-Gallés (2001). Testing learners at two proficiency levels (intermediate and advanced) allowed us to track a developmental sequence. A one-to-one relationship between category distinction and lexical contrast would be visible, if improved categorization enhanced lexical decisions. (Strictly speaking, even if this is found, the causality relation is still not demonstrated.) However, if the acquisition of a lexical contrast is not accompanied by a change in categorization, then this suggests that category formation is not a pre-requisite for the acquisition of a lexical contrast.

III On the acquisition of the /y/-/u/ and /œ/-/ɔ/ phonological contrasts

Here we introduce a model of phonological acquisition, DMAP (*Direct Mapping from Acoustics to Phonology*) that highlights a different flow of dependencies in acquisition. DMAP is captured by the four propositions stated in (1).

(1) *Direct Mapping from Acoustics to Phonology* (*DMAP*)

(a) L2 learners detect more acoustic cues in the raw percepts than what they use to perform a segmental categorization response.

(b) Detected features trigger revisions of the Interlanguage feature hierarchy in accordance with economy principles.

(c) Phonological lexical representations consist of feature matrices dependent on the Interlanguage feature hierarchy at the time of encoding.

(d) Minimal changes in phonetic category definitions triggered by phonological contrast obey economy considerations at this level.

In DMAP, the first step of the learning process resides in cue-based feature detection from the raw percepts (1a). The onset of new phone acquisition lies with the restructuring of the feature system guided by economy principles (1b). The encoding of new lexical contrasts involves phonological matrices enabled by revisions to the interlanguage feature hierarchy (1c). Economy requires the smallest modifications of previous phonetic values to reflect phonological contrast (1d). The encoding of lexical contrast is thus independent of and hence can precede reliable category formation.

According to DMAP (1a), adult L1-English L2-French learners can detect correlates of phonological features in the raw percepts of the input, and extract the relevant features, following Dresher and Kaye's (1990) cue-based learning. The assumption that feature detection is required for acquisition is not particularly controversial. However, it is necessary to recognize that the lack of robust discrimination response in the face of category assimilation in particular tasks does not mean that features relevant to the contrast cannot be detected, and therefore, it cannot be equated with auditory insensitivity. In other words, even though everything can be detected, not everything will be meaningful in terms of the L1 segmental categorization response. The acquisition of front rounded vowels by (US-)English-French learners offers a highly suitable case for examining the relationship between categorization and lexical representation of a phonological contrast. Levy and Strange (2008) showed that (US-) English-French learners initially experience perceptual problems within particular consonantal contexts in their categorization of front rounded vowels. These context-specific difficulties are (largely) overcome in advanced learners, although categories for such vowels are clearly not target-like given the persistence of categorization errors. The acquisition of /y/-/u/ and /œ/-/ɔ/ contrasts in lexical representations requires the detection of complex acoustic cues relevant to the features [back], [front], [high], and [round] (Fant, 1969). For instance, for a non-low vowel, the proximity of F1 and F2 can be an acoustic correlate of the feature [back] enhanced by lip rounding (Keyser and Stevens, 1994).

According to DMAP (1b), the perceptual system will detect correlates of {[front], [round]} combinations in French vowels but the phonological grammar initially fails to license such feature combinations, which are therefore ignored in lexical encoding at this stage. Thus, the rounded vowels are re-interpreted as back vowels by the L1-based interlanguage phonology, yielding the merger between target /u/ and /y/ in interlanguage, for example, which can also be subject to effects of contexts, as shown in Levy and Strange (2008). However, the detection of {[front], [round]} combinations in non-coronal contexts (as in *pube* [pyb] 'publicity' or *pub* [pœb] 'pub') means that the value [front] cannot be derived from context. Licensing failure triggers phonological acquisition: Repeated occurrences of {[front], [round]} combinations highlight the distinctiveness of the feature [round]. Thus, with Clements (2001: 71), we assume that "[f]eatures, nodes or tiers that are not employed in a given language remain *latent* in the sense that they remain potentially available, and may subsequently become distinctive or active as a result of language contact, internal historical change, and other dynamic factors influencing language development."

According to DMAP (1b), general principles of economy governing representations are also at play in revisions to the phonological state. This includes representational economy at the segmental level, but also the maximal use of a distinctive feature in the specification of inventories. Clements (2003, 2009) calls this optimization 'feature economy'. Symmetry is favored as a result of such maximal use of features. The contrastiveness of [round] means establishing {[front], [round]} and {[back], [round]} in vowel matrices as a symmetric reflex of feature economy. Indeed, the need to quickly compute a phonological representation of the input is best served if representations are constrained by economy. Clements (2001: 71–72) argues: 'phonological representations should be freed of superfluous representational elements, leaving only those that are essential to an understanding of lexical, phonological, and phonetic generalizations.' Not all feature values need to be specified to establish contrasts, and so they are not (Archangeli 1988; Avery and Rice 1989). Thus, if the classical feature system for vowel height is on the right track, high rounded vowels require a specification [+high] under the vowel height node (since [-high] is either a default value or simply absent), but mid rounded vowels do not. The reassembly of features into matrices for phonemes has a cost: /y/ might thus be acquired later than /œ/ since the cost of establishing a feature matrix for /y/ is expected to be higher than the cost of establishing one for /œ/, given the fuller specifications required for high vowels. In DMAP, therefore, upon exposure to L2, learner's phonological acquisition beyond the L1-induced initial state is mediated by general conditions on feature systems. Reflexes of economy are found at every level in L2 phonological systems (Altenberg and Vago, 1983; Broselow, Chen and Wang, 1998; Carlisle, 1998; Eckman, 1984; Halicki, 2010).

According to DMAP (1c), L1-English L2-French lexical representations involve only those feature matrices licensed by the interlanguage feature system at the time of encoding. Initially, target-language contrasts are merged, leading to spurious homophony. As the {[front]/[back] + [round]} matrices are acquired, rounded vowel contrasts can be lexically encoded. In the general case, DMAP does not guarantee that interlanguage lexical contrasts are represented by the same feature combinations across groups of learners with different L1s and at different proficient levels, in view of the L1-based initial state and amount of exposure.

According to DMAP (1d), phonetic category definitions must reflect phonological feature contrasts. However, the requirement of distinct category definitions does not require attunement to target-like category boundaries. Target-like boundaries require myriad adjustments. Hence, lexical contrasts can be established in advance of target-like phone values.

DMAP does not deny the role of categorization in processing; rather, it highlights what is strictly necessary for phonological acquisition. This crucially requires only the detection of acoustic correlates of phonological features in the raw percepts, not the complete overcoming of category assimilation. The establishment of interlanguage inventories occurs at two disjoint levels: the development of phonological feature matrices and the adjustments of phonetic category definitions (Maye, 2000; Maye, Werker and Gerken, 2002). Lexical encoding of a phonological contrast is expected to be largely independent of the attunement of phonetic categories to the L2 input. Despite an established lexical contrast, target-like categories might still be invisible in a segmental categorization task such as ABX.

IV. The experimental paradigm

4. 1. Materials

ABX categorization. A suitable method that combines discrimination and identification, but does not require word identification is ABX or AXB, where a listener has to match through mental comparison a token X to either token A or B, indicating his or her answer by pressing a button labeled A or B. In ABX/AXB tasks, listeners have to generalize over changes in voice or acoustic details (for example, changes in F0 or speech rate, varied segmental contexts inducing coarticulatory variation) in order to perform the matching of a token X to the other tokens, thereby demonstrating a robust categorization pattern (Dupoux, Pallier, Kakehi and Mehler, 2001). Depending on how the design implements those variables, the task can be made to enforce a more acoustic response or a more phonetically/phonologically sensitive one (see Højen and Flege, 2006 for discussion). In our task, different female voices and tokens were used so that a response could not be given merely on the basis of an auditory comparison.

We created CVC non-word pairs contrasting the vowel pairs /y/-/u/ and /œ/-/ɔ/ and the control pair /i/-/ɛ/ in two different consonantal contexts. There were eight pairs of non-words in each of six conditions (for a total of 48 pairs): labial context for /y/-/u/, labial context for /œ/-/ɔ/, coronal context for /y/-/u/, coronal context for /œ/-/ɔ/, controls with /i/-/ɛ/ (four in labial context, four in coronal context), and an additional control condition where consonants were different and the vowels were the same (all /u/ or all /y/) across the different pairings. All items were non-words in French with one exception ([lɔt]), a low frequency word[1]. The context combinations C_C were the same for all test vowels for a given place of articulation, for example, all the labial contexts were the same for /y/-/u/ and /œ/-/ɔ/ pairs. For the control conditions (vowel /i/-/ɛ/ as well as consonant), the C_C consonant combinations used were different from those used in the test conditions. Stimuli were recorded several times by two female French native speakers in a sound-isolated recording booth at a sampling rate of 44100 Hz with a 16-bit resolution, on a mono channel. Recordings were normalized for amplitude and spliced into separate sound files. Two renditions from each speaker were obtained for each non-word. One voice was used for the X token, whereas the other was used for the two different A and B tokens.

Stimuli were arranged in four different pairings for each pair: ABA, ABB, BAA, and BAB. For each pairing, the first two non-words were in one voice (e.g. AB), the third was in the other (X). The sound tokens used for ABA and ABB were different from those used for BAA and BAB. This yielded a total of 192 trials (one trial being a sequence of three non-words). The randomization was set such that the same pair in both "minimal pairings" ABB and ABA, for example, would not occur in the same block. Otherwise, all items were automatically randomized by the program for presentation to participants into six blocks. Blocks were separated by pauses. The inter-stimulus interval was 500 ms, and the response time-out was 2000 ms.

Lexical Decision with repetition priming. This experiment was designed closely following the method used by Pallier, Colomé and Sebastián-Gallés (2001). We selected four contrasts for the test: high vowels /i/-/y/ and /y/-/u/, and mid vowels /ɛ/-/œ/ and /œ/-/ɔ/ (see Appendix). In order for the comparison of ABX and lexical decision results to make sense, we focus our report on data obtained with the vowel contrasts that were included in both experiments: high vowels /y/-/u/ and mid vowels /œ/-/ɔ/ (along with the contrast /i/-/y/, which we call the "control condition" since US-English learners of French soon recover from this initial perceptual assimilation; Levy and Strange, 2008).

The stimuli were French words and pseudo-words. As much as possible, we avoided French pseudo-words that were reminiscent of English words. Forty words forming 20 minimal pairs based on the four contrasts were included. In addition, 40 French pseudo-words were created that formed 20 minimal pairs following the same pattern as the preceding words. Finally, 120 words and pseudo-words were also included to serve as filler items. Sixty were repeated in order to model the repetition pattern in place for the test words and pseudo-words, yielding a total of 180 filler items. Stimuli were recorded several times by a female French native speaker in a sound-isolated recording booth at a sampling rate of 44100 Hz with a 16 bit resolution, on a mono channel. Two different sound tokens were selected for each item.

Due to the need to find common minimal pairs contrasting those sounds, it was difficult to match the words in terms of frequency exactly. However, the word pairs containing the contrasts /y/-/u/ and /œ/-/ɔ/ were in the aggregate closely matched in frequency, as measured by averaging the frequency – written or spoken – of the words used for a given contrast. The /i/-/y/ control condition contained words that were overall lower in frequency. The verification was performed using Lexique 3.70 (New et al., 2001, 2004), revealing the following distribution:

Insert Table 1 here

The verification of frequencies across contrasts was performed on all word forms (listed in Appendix 1) as well as their homophones (for example, the frequency of [pɔʁ] was combined across that for *port* 'harbour', *porc* 'pork' and all other homophones that share the form [pɔʁ]). Since a listener can activate any of the homophones during the task, there was no written presentation of [pɔʁ]). Most words had more than one homophone.

Even though the /y/-/u/ and /œ/-/ɔ/ tokens were matched in frequency, it is unclear how the French frequency of those words translates into L2 learners' familiarity with those words. We therefore administered a familiarity questionnaire (using the most frequent of all homophones for each word form; details and results are presented in section 4.4) to the intermediate group at the end of the experiment. Informal debriefing revealed that all words were known by the advanced learners.

Four counterbalanced lists of 260 stimuli were created in the following way: In each list, one member of each minimal pair appeared (e.g. /din/ from /din/-/dyn/) and was followed, 8 to 20 items further down in the list, either by the other item in the minimal pair (e.g. /dyn/), or by itself (e.g. /din/). The members of a given minimal pair were counterbalanced across the lists. Different sound tokens were used for the repetitions, so that none was actually heard twice. The inter-stimulus interval was 2500 ms, and the response time-out was 2200 ms.

4.2. Participants

Three groups were tested on both experiments: advanced and intermediate English-French learners and native speakers of French. A group of native English speakers with no exposure to French, or any language with front rounded vowels, also completed the ABX task. They did not complete the lexical decision task because it requires knowledge of French.

The intermediate learners (n = 38, 9 males) are native speakers of English. They all started to learn French at or after the age of 10 at school. Learners' proficiency groups were determined on the basis of current course enrollment. Intermediate learners were in their fourth or fifth semester of college French at a major US university. (Magnan (1986) has shown that most students from such classes are at the intermediate-high level on the ACTFL scale, as measured by the ACTFL Oral Proficiency Interview.) Their mean age was 20.2 years (range 18-36). Their instruction in French at the university ranged from 1 to 6 (average 3.8) semesters of French. None of them had spent a large amount of time in a French-speaking country (on average 1.5 weeks, ranging between 0 and 7 weeks), except one participant who had spent 6 months in France. Outside of class, they all had none to very little regular exposure to written French (on average, 1.2 hours/week) or to spoken French (on average 50 minutes/week). All of them were native speakers of (US-) English, and none of them grew up bilingually or were highly proficient in another language containing the relevant contrasts, even though some had some knowledge of additional languages. Of these additional languages, the only one with /y/-/u/ and /œ/-/ɔ/ contrasts was German, low proficiency in which was reported by five participants. All of them had normal hearing.

Advanced learners (n = 20, 10 males) were graduate students or French instructors at the same university. They all started to learn French at or after the age of 10, with the exception of one who started in grade 3 at the age of 8. Their mean age was 30 years (range 22-49). All had spent some time in at least one French-speaking country, ranging from five months to three years and longer (range 20–156 weeks), with the exception of one (who had spent seven weeks in France). As confirmed by one of the experimenters, who is a native speaker of French, all participants reported a high to very high level of general proficiency in French, even though some of them retained a strong foreign accent. All of them had daily sustained to extensive exposure to spoken French (average 12.3 hours /week) and/or written French (average 7 hours/week). All of them are native speakers of (US-) English; none of them grew up bilingually. Most of our advanced learners had some knowledge of one or more additional languages, but none had had any early exposure to those languages: German (10), Spanish (9), Italian (8), Arabic (1), Breton (1), Catalan (1), Chichewa (1), Chinese (1), Dutch (1), Greek (1), Haitian Creole (1), Hebrew (1), Ojibwe (1), Picard (1), Portuguese (1), Russian (1) Serbian/Croatian (1), Welsh (1). All of them had normal hearing.

French native speakers (n = 10, 1 male) were either faculty or graduate students at the same university at the time of the study. All of them were exposed to French daily, due to personal and/or professional reasons. Their mean age was 28 (range 24-33) years. All of them had normal hearing.

We also tested a control group of 13 (2 males) naive (US-) English native speakers (‘English monolinguals’), each of whom had had no significant exposure to Dutch, German, French, any Scandinavian or Chinese language, Korean, or Finnish. This group only took part in the ABX task. They were undergraduate students at the time of testing; their mean age was 19 (range 18-21) years. All of them had normal hearing.

4. 3. Procedures

A list of the test words used in the lexical decision task was distributed to the teachers of classes from which we recruited the intermediate learners several weeks before the experiments. Students were not informed that those words would be part of a later experiment. Teachers were told to try to use the words in class or in assignments, but they were unaware of the purpose of the experiments. Exposing students to those words beforehand was done to reduce the number of exclusions due to high error rates.

All participants were tested individually in a quiet room. Both experiments were administered in a single session. After completing a linguistic background questionnaire, participants first took part in the lexical decision experiment, and then continued with the ABX task (except for the control group of English monolinguals who took part only in the ABX task); at the end, intermediate learners were given a list of words and asked to indicate which ones they knew.

All test words used in the lexical decision experiment were included in the list. The presentation of the stimuli was fully controlled by Dell personal computers. Auditory stimuli were presented through Sennheiser HD515 headphones. The experimental presentation was controlled by the DMASTR software (DMDX) developed at Monash University and at the University of Arizona by K.I. Forster and J.C. Forster.

ABX categorization. Listeners heard three non-words in a row, A, B and X, and were asked to decide whether X is like A or like B. The experiment was preceded by a short training session of eight trials with feedback. The goal beyond simple task familiarization was also to speed up reaction times to ensure more automatic responses and minimize strategic responding. The total duration of the experiment was around 20 minutes.

Lexical Decision with repetition priming. Participants were instructed to decide as fast as they could for each item whether or not it is a real French word. A short training session of eight items with feedback familiarized them with the experiment and the pace. Each participant was randomly assigned to one of the four lists. The total duration of this experiment was around 20 minutes.

4.4. Acoustic analysis and lexical familiarity

In order to ensure that stimuli are comparable in both experiments, we performed acoustic measurements of F1, F2 and F3 for each token. Results for F1 and F2 are presented in Figure 1.

Insert Figure 1 about here

Figure 1 demonstrates the comparability of the two sets of stimuli. As shown, mid vowels are closer in the acoustic space in terms of F2 than high vowels. The direct comparison of average formant values across both experiments is however not very meaningful because the ABX items are produced by two different speakers; these are also on average three times more numerous than lexical decision tokens, so that more variability is visible in ABX tokens. F1, F2 and F3 values differ statistically for all four vowels, being lower in the lexical decision tokens, with the exception of F1 for /œ/. This difference is mainly due to the voice characteristics of the second ABX speaker (for instance, her F2 is systematically higher than the first speaker's F2). For our purposes, it is important to note that even if there is some systematic difference in the spectral characteristics of our ABX vs. lexical decision stimuli, this difference affects all vowels equally. Observing a better performance on one vowel contrast rather than another will therefore not be easily attributable to acoustic characteristics of the stimuli only.

Duration, however, can be more clearly affected by consonantal context, and may play a role in facilitating discrimination. In particular, our two stimuli sets differ in one crucial dimension: due to considerations of word frequency and familiarity for learners, all but one pair of lexical decision words in the /œ/-/ɔ/ contrast end with a rhotic consonant (a voiced fricative [ʀ] or a uvular [ʁ]). This is not the case in ABX stimuli for the same contrast, where there is a more balanced proportion of lengthening vs. non-lengthening consonants (voiced fricatives vs. obstruents). It is also not the case for the /y/-/u/ contrast. Vowel durations for all stimuli are presented in Table 2.

Insert Table 2 about here

Table 2 shows that, for /y/-/u/, duration is the same in lexical decision and ABX tokens. For /œ/-/ɔ/, however, duration (all items) is longer in lexical decision, possibly because of the rhotic context in the words for this contrast. We separately measured ABX tokens which also end in a lengthening consonant (voiced fricatives). When comparing these tokens (L-only) with the lexical decision items, this difference disappears.

The word familiarity questionnaire given to the intermediate learners revealed that the intermediate learners knew on average 71% of the words. Considering only those participants who were included in subsequent analyses (see Section V for details), average familiarity increased to 74%. Only a few words were unknown to almost all of our participants. Overall, word familiarity was 60% for /i/-/y/ contrast pairs, 61% for /y/-/u/ pairs, and 85% for /œ/-/ɔ/ pairs. We will return to the potential significance of word familiarity and acoustic analysis data for the lexical decision results in the general discussion section.

V. Results

5.1. Screening

Given that reaction times are the dependent measure and since RTs are computed over correct answers, it was imperative to exclude subjects with too high an error rate. Following the practice of White, Melhorn and Mattys (2009), we excluded participants who had lexical decision accuracy rates below 75% across all trials from the analyses. This resulted in the elimination of results from 18 of the 38 intermediate learners, but from none of the advanced learners and French natives. For the ABX task, a total of three participants had to be excluded: one intermediate and one French native speaker were excluded because they apparently did not follow instructions, reversing responses systematically (they pressed B for A and vice-versa). Another French native speaker was excluded because of high error rates on the filler condition (higher than 1.5 SD from the mean). No advanced learner was excluded; all participants excluded from ABX task were also removed from the lexical decision task. The final number of participants is 19 intermediates, 19 advanced learners, and 8 French native speakers. None of the participants in the control group of English monolinguals, tested only on ABX, had to be removed (n = 13).

5. 2. ABX Categorization

Error rates were calculated across both segmental contexts first and are displayed as a function of group and vowel pair in Figure 2. The results of the

ABX categorical discrimination task reveal that the intermediate and advanced learners performed similarly to each other on all contrasts. Both groups performed more accurately on the /y/-/u/ contrast than on the /œ/-/ɔ/ contrast.

Insert Figure 2 about here

Error rates were used as the dependant variable in an Analysis of Variance (ANOVA) with group (advanced, intermediate, French native-speakers and English monolinguals, between subjects) and vowel pair (‘control’/i/-/ɛ/, /y/-/u/ and /œ/-/ɔ/ contrasts, within subject) as declared factors. We observed a main effect of group, $F(3, 55) = 7.6$, $p < .001$, and vowel pair, $F(2, 110) = 222.5$, $p < .0001$, as well as a significant interaction between the two, $F(6, 110) = 2.3$, $p <. 05$. Subsequent analyses, setting aside the group of English monolinguals, restricted to each condition and declaring the factor group revealed that there was no group effect on the control contrast, $F(2, 43) = 0.1$, $p > .1$, but a significant group effect on the other two vowel pairs: For the /y/-/u/ contrast, $F(2, 43) = 3.2$, $p <. 05$, and for the /œ/-/ɔ/ contrast, $F(2, 43) = 5.1$, $p < .05$. The same analysis with the group ‘English monolinguals’ had shown that the group factor was significant, $F(3,55) = 4.6$, $p < .01$, also in the control contrast (due to the slightly higher error rate of this group).

These results suggest that the learner groups and the French natives behaved similarly on the control vowel pair, but their error rate pattern was different for the /y/-/u/ and /œ/-/ɔ/ pairs. A visual inspection of Figure 2, however, suggests that this effect could be mainly due to the French native speaker group behaving differently, with fewer errors than all other groups. This calls for an analysis restricted to both learner groups. An ANOVA declaring the factors group (intermediate, advanced) and vowel pair (/ɛ/-/i/, /y/-/u/ and /œ/-/ɔ/ contrasts) showed a significant effect of vowel pair, $F(2,72) = 184.6$, $p < .0001$, but crucially, no significant effect of group, $F(1, 36) = 1.6$, $p > .2$. The interaction was now absent, $F(2, 72) = 0.5$, $p > .5$, suggesting that the interaction reported above was due to the native speaker group. This confirms the impression that both learner groups behaved in the same way on each vowel pair. Separate two-sample *t*-tests confirmed that the intermediate and advanced groups were not significantly different on any vowel pair ($p > .1$ in all cases). So far, results show that both learner groups behave alike despite the difference in their exposure to French. Advanced learners produced slightly fewer errors overall than the intermediate group, but they were never significantly better. Both learner groups differed from the native speakers on all experimental vowel pairs, but not on the controls. The comparison with the group of English monolinguals on the /y/-/u/ contrast showed that English monolinguals differed from the French native speakers, $t(19) = 3.4$, $p < .01$, from the advanced, $t(30) = 3.0$, $p <.01$, and from the intermediate learners, $t(30) = 2.1$, $p < .05$. On the /œ/-/ɔ/ contrast, they differed significantly from the French natives, $t(19) = 3.3$, $p <.01$, and the advanced $t(30) = 2.2$, $p < .05$, but not from the intermediate learners, $t(30) = 1.5$, $p >. 05$.

A further analysis of error rates according to the segmental context revealed the following pattern (see Table 3).

Insert Table 3 here

Native speakers did not show any significant context effect, but intermediate learners were more accurate in the labial context for both contrasts (see Levy and Strange, 2008 for similar results). Advanced learners have overcome the context effect for mid vowels, but still had a marginally significant effect for the high vowel contrast. Additionally, we analyzed items containing a lengthening coda (voiced fricatives) separately from the others. Error rates in each condition are presented in Table 4. If longer durations are clearly involved in facilitating categorization, error rate is expected to differ in lengthening (L) vs. non-lengthening (NL) conditions.

Insert Table 4 about here

A global mixed model analysis of variance on error rates declaring the factors group (advanced, intermediate, French native-speakers and English monolinguals), vowel height (mid vs. high), and lengthening context (L vs. NL) showed a significant main effect of group, $F(3,220) = 16.4$, $p <.0001$, and vowel height, $F(1, 220) = 94.4$, $p <.0001$, but no effect of lengthening context, $F(1,220) = .3$, $p >.5$. No interaction was significant. Similarly, in the group-specific analysis, the effect of lengthening context did not affect performance in any group (all $F <1$, all $p >.3$). These data suggest that acoustic cues that may be perceived more clearly through longer durations do not interact with categorization performance during the ABX task.

5. 3. Lexical decision with repetition priming

Lexical decision latencies were measured from the onset of each word. Incorrect responses to word targets were not analyzed. Reaction times for correct responses above 2200ms or below 300ms (computed separately for each participant's target and baseline trials) and any RTs more than 2.0 S.D. units away from the overall mean RT for each subject were trimmed by setting it equal to the cutoff value (2200ms or 300ms, respectively). This occurred for 6.3% of the data in the intermediate group, and for 4% of the data in the advanced and native speaker groups combined. Results are presented for words first, and then for non-words. Reaction times are reported separately for each group and each contrast in Table 5. For each group and contrast, there are four latencies: two for the *repetition* condition (rep), and two for the *minimal-pair* condition (mp). The latencies correspond to the order of presentation of forms. In the repetition condition, the forms are occurrences of the same word. In the minimal-pair condition, the forms contrast minimally. Priming is measured by subtracting the second latency from the first.

Insert Table 5 about here

A shorter decision time for the second occurrence is expected if the word is a repetition across both occurrences. In the minimal-pair condition, if the word in second position is experienced as different, such facilitation is not expected. However, if learners encoded members of the minimal pair as target-deviant homophones, the presentation of the first member will pre-activate the second member, resulting in faster lexical decision for the occurrence of the second member of the pair, in a manner parallel to the repetition condition. In Figures 3 to 5, bars represent the priming effect, i.e. the amount of facilitation in reaction times with which participants correctly answered for the words in each condition. The white bar corresponds to priming obtained for a repetition; the dark bar corresponds to priming obtained with a minimal pair. Latencies to the first vs. the second occurrence of a word pair are compared within each group with an analysis of variance using a linear mixed-model. RT is the dependent variable and there are three factors: “contrast” (all four contrasts were included in the model, but we report on the three contrasts /i/-/y/, /y/-/u/ and /œ/-/ɔ/ only), “condition” (minimal pair, repetition) and “time” (presentation time 1, presentation time 2).

In the lexical decision with repetition priming task, results show that all participants showed facilitations (positive priming) in response times in the repetition condition (white bars). Crucially, in the minimal pair condition, neither advanced learners nor native speakers showed any significant priming for either the /y/-/u/ or the /œ/-/ɔ/ minimal pairs. This constitutes evidence that they are not treating the minimal pairs as homophones. In stark contrast, intermediate learners produced response-time facilitations for the /y/-/u/ minimal pairs (dark bar), but the priming for the /œ/-/ɔ/ minimal pairs does not reach statistical significance. In general, the native speaker group had shorter reaction times than the learners. Within each group, there is also a tendency for the words in the /œ/-/ɔ/ contrast to yield faster reaction times. As pointed out by a reviewer, faster reaction times could make it harder to obtain a significant priming effect in the minimal pair condition for this contrast. However, despite the faster reaction times, learners produced priming in the repetition condition. Faster reaction times do not seem to prevent priming. The statistical analysis reveals that in all groups, the interaction between contrast and condition is not significant. The reaction times obtained in the different conditions do not vary with contrast. In other words, the participants responded to each contrast with comparable latencies in each condition. We now consider each group in turn.

Insert Figure 3 here (native speakers)

The native speaker group shows a significant effect of time ($F[1, 105] = 14.6$, $p < .001$), of condition ($F[1, 105] = 8.11$, $p < .005$), as well as a significant interaction between time and condition ($F[1, 105] = 6.88$, $p < .01$). Other effects and interactions did not reach significance. The significant interaction between time and condition suggests that the latency difference due to time is not the same in each condition. Planned comparisons reveal that the priming effect, i.e. the difference in RT due to time (time 1 – time 2) is indeed significant only for the repetition condition ($F[1, 105] = 20.8$, $p < .001$), not for the minimal pair condition ($F[1, 105] = .730$, $p > .3$). The lack of interaction time*contrast suggests that the RT-difference due to time is comparable for each contrast. A detailed examination of each contrast reveals that this pattern is found in each case: significant priming in the repetition condition, but no priming in the minimal pair condition (in one contrast, /i/-/y/, priming in the repetition condition is marginal ($F[1, 105] = 3.6$, $p < .060$), which may be due to the small sample size of this group).

Insert Figure 4 here (advanced)

In the case of the advanced learners, a similar pattern is observed. There is again a significant effect of time ($F[1, 270] = 23.6$, $p < .001$), of condition ($F[1, 270] = 19.2$, $p < .001$), as well as a significant interaction between time and condition ($F[1, 270] = 11.5$, $p < .001$). In addition, there is a main effect of contrast ($F[3, 270] = 2.8$, $p < .039$) and a significant time*contrast interaction ($F[3, 270] = 2.9$, $p < .031$), which suggests that not all contrasts behave similarly in terms of the time-related latency difference. Figure 4 reveals that this significant interaction may stem from the large negative priming obtained for the control contrast. Importantly, for both the /y/-/u/ and /œ/-/ɔ/ test contrasts, the advanced learners displayed a behavior similar to that of the native speakers. Planned comparisons confirm that the priming effect is significant only for the repetition condition ($F[1, 270] = 34.1$, $p < .001$), not for the minimal pair condition ($F[1, 270] = 1.0$, $p > .3$). A detailed examination of each contrast reveals that this pattern is found in each case. As in the case of native speakers, there is one contrast (/œ/-/ɔ/) for which the statistical value in the repetition condition is marginal ($F[1, 170] = 2.9$, $p < .09$).

Insert Figure 5 here (intermediates)

For Intermediate learners, there is a significant effect of time ($F[1, 270] = 27.7$, $p < .001$), of condition ($F[1, 270] = 13.7$, $p < .001$), as well as a significant interaction between time and condition ($F[1, 270] = 7.7$, $p < .006$). In addition, there is a main effect of contrast ($F[3, 270] = 6.0$, $p < .001$) and a significant time * contrast interaction ($F[3, 270] = 4.1$, $p < .007$), which suggests that not all contrasts behave similarly with respect to the time of presentation. Strikingly, for the intermediate group there is also a significant triple interaction time*contrast*condition ($F[3, 270] = 3.3$, $p < .020$). This suggests that the time difference between time 1 and time 2 varies in each condition as a function of contrast. Planned comparisons confirm that the priming effect is significant for the repetition condition ($F[1, 270] = 32.3$, $p < .001$) as in the two other groups. For the minimal pair condition, unlike in the other groups, the difference between time 1 and time 2 is marginal ($F[1, 270] = 3.0$, $p < .08$). This finding together with the triple interaction warrants a closer examination. Figure 5 clearly shows that the intermediate learners differ from both other groups on the test contrast /y/-/u/. As confirmed by the contrast-specific analysis, there is significant priming for all three contrasts in the repetition condition (/i/-/y/: $F[1, 270] = 10.3$, $p <$

.001; /y/-/u/: $F[1, 270] = 10.9$, $p < .001$; /œ/-/ɔ/: $F[1, 270] = 4.04$, $p < .045$). For the minimal pair condition, there is priming neither for /i/-/y/ ($F[1, 270] = .07$, $p > .7$) nor for /œ/-/ɔ/ ($F[1, 270] = 1.1$, $p > .2$). For /y/-/u/ in the minimal pair condition, there is a highly significant priming ($F[1, 270] = 18.4$, $p < .0001$).

There was no significant facilitation on any non-word condition (repetition or minimal pair for all three contrasts) for the native speakers (all $p > .1$) and for the intermediate learners (all $p > .1$); for the advanced learners, no priming was observed on any condition (all $p > .05$) with two exceptions: in the repetition condition for /i/-/y/ and in the minimal pair condition for /y/-/u/ (both $p < .02$).

Let us summarize our main findings for this experiment. Intermediate learners exhibited priming effects indicative of spurious homophony on the /y/-/u/ contrast, but produced no facilitations across minimal pairs on the /œ/-/ɔ/ contrast. Advanced learners patterned like the native speakers on all three contrasts; their reaction times were comparable to the intermediate learners.

VI. General Discussion

Our empirical findings bring up a curious anomaly for standard assumptions according to which the development of new categories is a necessary prerequisite for lexical contrast. Our advanced learners had established lexical contrasts based on all tested French minimal pairs (/i/-/y/, /y/-/u/, and /œ/-/ɔ/), but exhibited persistent perceptual errors in the categorization of contrasts. For intermediate learners, the picture is even more intriguing. On the lexical decision task with repetition priming, these learners displayed spurious homophony showing merger of contrasts between minimal pairs for the high vowels /y/-/u/, but showing lexical contrast for minimal pairs involving mid-vowels /œ/-/ɔ/ in French. However, on the ABX task, these same learners exhibited a significantly higher average error rate for the /œ/-/ɔ/ contrast than for the /y/-/u/ contrast (37% vs. 15%). It seems that the establishment of a lexical contrast is independent of the previous acquisition of phonetic categories as observed in categorization tasks.

Our results on the /y/-/u/ contrast in intermediate learners confirm previous findings of spurious homophony in L2 learners – attributed to the absence of well-defined categories (e.g. Pallier et al., 2001). Yet, our results on the /y/-/u/ contrast in advanced learners also show for the first time that spurious homophony can be resolved with more experience.

Hence, we observed specific patterns of breakdown and recovery in lexical representations on the /y/-/u/ contrast in the acquisition of French with no benefit for categorization. In addition, we observed evidence of lexical representations on the /œ/-/ɔ/ contrast despite significant categorization errors. These observations invite a reconsideration of the causal link between categories and lexical contrast in second language acquisition. Our advanced learners encoded lexical contrasts even though their categorization performance was not different from that of intermediate learners. If one assumes that robust categorization of a phonological contrast is required for lexical encoding of this contrast, our results are puzzling.

The similarity of the two learner groups on the ABX tasks (despite significant differences on the lexical decision task) supports the conclusion that categorization and lexical encoding are separately acquired. In fact, the only difference between both groups was related to the stability of the categorization, as seen through resistance to context effects. The intermediate learners (but neither the advanced learners nor the French natives) experienced a significant context effect during categorization (p <.05, with more errors in the coronal context). Context effects typically arise when a perceived phonetic value for a particular segment is attributed to (originating from) the surrounding segmental context, triggering a different categorization of the segment. Here, the perceived front value of the vowel (/y, œ/) is treated as coarticulation emanating from its coronal context; in English, a back vowel is fronted through coarticulation with coronal consonants, so that listeners attribute frontness to coronal coarticulation. As a result, values of /y, œ/ are interpreted as instances of /u, ɔ/, creating a case of single-category assimilation mediated by context.

Learners also experienced greater categorization difficulties with mid vowels /œ/-/ɔ/. This categorization problem presents us with a paradox. The mid range of the learner's L1 perceptual map is dense. Indeed, Levy (2009a) shows that French /œ/ is assimilated to English vowel categories /ɚ/, /ʌ/, /ʊ/, /ɔ/ and /u/ in a pattern affected by context and experience. This several-category assimilation pattern, by reducing overlap, might reasonably lead to the expectation of fewer errors in the ABX. This, however, was not observed, presumably as a result of processing load. Indeed, from the point of view of the Interlanguage categorization of target language phones, (partially overlapping) L1 categorization responses must be suppressed. The more there is to suppress, the greater the cost. Hence, the richness of the perceptual map in the mid field would feasibly increase the computational load associated with this categorization, and trigger the higher error rate we observed.

Conclusions about the nature of L2 sound systems from these observations require a careful characterization of the evidence, as it might be compatible with a range of scenarios. The degree to which these asymmetries highlight a development sustained by phonological computations as discussed in DMAP can be weighed against the degree to which these asymmetries may be due to aspects of the stimuli: frequency, familiarity, as well as orthographic and acoustic cues. The discussion focuses on asymmetries in priming for /œ/-/ɔ/ versus /y/-/u/ contrasts in the lexical decision task for intermediate learners (only).

The study was designed so that the frequency analysis of the words representing each contrast showed that the words for both /y/-/u/ and /œ/-/ɔ/ were comparable. However, the frequency of the sounds themselves could also differently influence the lexical encoding of words that contain those sounds. For instance, encoding words with /œ/ could be made easier because of the frequent use of this sound in the language in general. In French, the sound /œ/ is highly frequent (Hume and Bromberg, 2005). It occurs both as a realization of schwa in function words such as *je* 'I', *te* 'you-acc' and *me* 'me-acc' and *le* 'the, it-acc', and as an independent phoneme /œ/. French schwa systematically deletes or undergoes other readjustment in closed syllables, unlike the phonemic /œ/ that we examined. It is clear that the two uses of [œ] must be distinguished lexically. In the Brulex corpus, Hume and Bromberg (2005) establish similar frequencies for the two phonemic mid front rounded vowels (/œ/ and/;/ -Log$_2$ probability: 6.924) and the high front rounded /y/ (-Log$_2$ probability: 6.248), with /y/ slightly more frequent. For phonemic /œ/ alone, the type used in our experiment, frequency is much lower (-Log$_2$ probability: 8.525). Similar frequencies of words and vowels in those words make the plausibility of a frequency explanation rather remote.

L2ers' familiarity with the vocabulary might offer a more promising alternative. Higher familiarity with the test words in one contrast over another might have facilitated the encoding of a lexical contrast for those words. After completion of the experiment, we asked the intermediate learners to fill out a questionnaire about how familiar they were with the words used in the experiment. The words in the /œ/-/ɔ/ set were more familiar to the learners than the words in the /y/-/u/ set or in the /i/-/y/ set. The average familiarity for the /œ/-/ɔ/ word pairings was 85%. The average familiarity for the /y/-/u/ word pairings was 61%. There are two kinds of evidence that suggest that familiarity is not a sufficient explanation for the results reported here. To examine whether familiarity is driving the presence of repetition priming for the /y/-/u/ contrast (and the absence of it for /œ/-/ɔ/), we declared familiarity as a covariate in an ANOVA performed on latencies for the intermediate group, restricted to the three contrasts /i/-/y/, /y/-/u/ and /œ/-/ɔ/. The main effect of familiarity was not significant ($F[1, 213] = .5$, $p > .4$). The main effect of contrast is now marginal, $F[2, 207] = 2.3$, $p < .09$. All the values in the analysis by contrast were unaffected: the priming in the minimal pair condition for /y/-/u/ remained significant. The absence of priming for the /œ/-/ɔ/ contrast in the minimal pair condition was also confirmed. An analysis of the individual distribution of intermediate learners was performed to examine the potential effect of familiarity in greater detail. The group of intermediate learners was split into 2 groups, defined as the "high familiarity" group (n = 9, on average 81% of words are familiar to this group) vs. the "low familiarity" group (n = 10, on average, 69% of words are familiar to this group). These groups differed significantly in terms of familiarity ($t(17) = 6.03$, $p < .0001$), but not in terms of error rate (17.2 % and 17.9 % errors respectively, $t(17) = 0.4$, $p > .1$). Thus, familiarity scores seem dissociated from the lexical decision error rates. In sum, no significant effect of familiarity could be detected.

Orthographic knowledge might provide another alternative explanation for the encoding of a lexical contrast in advance of reliable categorization. Following this line of reasoning, Weber and Cutler (2004), Cutler, Weber and Otake (2006) and Escudero et al. (2008) proposed that L2 learners might directly deploy orthographic knowledge to acquire the contrast. Considering the number of alternate graphemes for an opposition, one can make the prediction that fewer grapheme options might help in establishing a contrast. In our case, the different graphemes for /œ/-/ɔ/ outnumber the graphemes used to represent /u/-/y/, seemingly predicting easier encoding for /y/-/u/. On the other hand, the graphemes used for /y/-/u/ might interfere more strongly with English grapheme-phoneme correspondences. In our case, /y/ and /u/ are both represented by distinct graphemes or grapheme combinations <u> and <ou>. The sounds /œ/ and /ɔ/ are represented as <eu> or <œu> versus <o> or <au> (e.g. for *Laure*). Following this possibility, different orthographic combinations for /o/ in French, <eau>, <o>, <ôt>, <ot>, <au> might lead to spurious minimal-pairs distinctions: <pot> /po/ 'pot' vs. <peau> /po/ 'skin'. To our knowledge, this phenomenon does not arise. It remains somewhat unclear whether the overlap with English orthography might interfere with the graphemes used for the /y/-/u/ contrast, since the phoneme /u/ corresponds to the grapheme <u> in English, interfering with the grapheme for /y/ in French (also <u>). The correspondence of <ou> to /u/ is not entirely new to English-speaking learners of French, since in English orthography <ou> occasionally corresponds to /u/, as in <you>, <coup>, and <mousse>. In sum, it is feasible that orthographic evidence may help learners focus resources on a particular contrast, speeding acquisition, but in our precise set of data, orthography makes no clear predictions of asymmetries in the representation of contrasts in intermediate learners for the test vocabulary.

Last but not least, we consider the possibility that this asymmetry results from the sound structure of the vocabulary across contrasts. The vocabulary items were selected on the basis of frequency exigencies and vocabulary knowledge within the limits of the sound pattern of French. As a result, all but one item in the /œ/-/ɔ/ minimal pairs contained a rhotic coda, realized as a (de)voiced fricative. This was not the case for /y/-/u/: only one minimal pair involved a rhotic coda. In French voiced fricatives (including uvular /r/) induce lengthening of immediately preceding vowels. Longer vowels could provide better formant cues, leading to the speedier establishment of rounded vowel contrasts. If longer vowels were the only explanation for the asymmetry in the lexical decision data, length should also affect the results of the ABX task. However, we found that /y/ and /u/ were shorter but better discriminated; furthermore, error rates were the same in lengthening and non-lengthening contexts. Vowel length does not affect categorization, and it is unlikely that differences in vowel length alone would facilitate the encoding of /œ/-/ɔ/ minimal pairs.

In contradistinction to this acoustic account, the words with a rhotic coda might allow a different phonological representation of this contrast in L1-English L2-French interlanguage. On this account, precisely in the context of a rhotic coda, intermediate learners are able to represent the minimal pairs with rounded mid vowels that contrast in the French lexicon on the basis of the feature [front] vs. [back] as minimal pairs contrasting on the basis of a central /ɚ/ or /ʌ/ vs. back /ɔ/. This is compatible with the category assimilation patterns for French mid round vowels found in Levy (2009a).

In sum, vowel length differences, orthographic cues, word familiarity and word frequencies do not satisfactorily account for our data set. The data strongly point to a phonological explanation.

VII. Contrast and discrimination of rounded vowels in DMAP

In view of the limitations of non-phonological approaches to these data, we now consider these data as support for DMAP. In DMAP, the learner's processing system at the outset of L2 acquisition includes a universal acoustic space defined by a general perceptual mechanism that extracts feature combinations from raw percepts, a phonological grammar that licenses the feature combinations extracted, and a L1 phonetic space reflecting the phonetic category definitions of the L1 phonological grammar on a continuum. The universal acoustic space enables the detection of {[front]/[back] + [round]} combinations in the raw percepts. Initially, these combinations are ignored in categorization and lexical encoding, as a result of phonological merger, in which detected (front) rounded vowels are corrected as back vowels, so that the acquisition of L2 vocabulary conforms to the L1-induced feature system, yielding spurious homophony. However, {[front]/[back] + [round]} combinations repeatedly detected in the French input trigger a change in the feature system. This change is driven by the need to parse the input with greater efficiency. This requires revising the feature system by (re)assembling matrices for the phone contrasts (in the spirit of Lardiere, 2009). Recovery from L2 initial-state phonological merger leads then to further development.

In DMAP, changes to the feature system obey general economy constraints on feature accessibility and feature (re)assembly. It is generally expected that phonemes that are underspecified in feature values will be acquired first, as a reflex of computational complexity. Setting the tense-lax/open-closed distinction aside, since it is mostly context-dependent in French (and not expected to play a significant role), /y/ requires a vowel height specification, whereas /œ/, being neither low nor high does not. If this is right, mid-vowel lexical contrasts might be established ahead of high-vowel lexical contrasts. Changes to the feature system must take place in the face of phonological processes. Thus, V-place instantiation of the features {[coronal] + [labial]} can be inherited contextually, so that L2 input can be reconciled with the interlangage phonology. Likewise, we theorized that words in which a rounded front mid vowel is followed by a rhotic coda could be phonologically reinterpreted as having a central (rhoticized) vowel. The rhotic element could be preserved on the vowel or assigned to the coda, leading initially to lexical contrasts involving central vs. back vowel specifications, rather than front vs. back ones. Thus, distinct phonological reasons can trigger the early establishment of lexical contrasts in advance of robust category distinctions based on frontness.

Before the opposition between front and back is identified as a feature of contrast for rounded vowels, rounded vowels are (generally) lexically encoded as back vowels by merger—creating spurious repetition priming. Once [labial] is assigned to the V-place (i.e. [round]+[front] is enabled), presumably as the result of input in which those features are not reducible to consonantal context (as would be the case for a coronal coda), contrasting feature matrices for front vs. back rounded vowels can be used in lexical representations. The lexical representations of previously established vocabulary entries presumably cannot be changed all at once to reflect the new matrices. Such a change would require that non-target-like lexical entries previously established include indexical information (e.g. /u*/) indicating that certain encodings of back vowels do not fully match the representations extracted from the input, as suggested by Hayes-Harb and Masuda (2008). Such a device is not strictly necessary for acquisition to take place, however. Indeed, target-like lexical representations can simply be acquired on the basis of positive evidence, once the phonological merger is overcome. There should be a lasting effect of the L2 initial-state phonological merger, even as new phonetic categories are established. A delay between what is phonologically possible at a given moment and what is lexically represented is to be expected.

Crucially, in DMAP, the acquisition of a lexical contrast for rounded vowels involves at its core the development of {[front]/[back] + [round]} feature matrices. Such matrices greatly underspecify the phonetic details of category definitions. As a result of under-specification of phonetic categories by the feature system, an inventory of phones can be described by a wide range of analyses (a point made manifest by phonological theory). In view of this, the presumed acquisition route from phonetic distinction to phonological contrast presents the learner with a central learnability problem: in principle, L2 inventories could be carved out in a large variety of ways. On a phonological approach, the acquisition of segmental inventories is guided by feature accessibility. DMAP offers an economy-driven mechanism that would obviate this learnability problem. Phonological contrasts require distinct category definitions but these do not need to be target-like. Economy dictates that the phonologically triggered change to category definitions should be the smallest change consistent with phonetic distinction. Indeed, although advanced learners clearly acquired lexical contrasts for rounded vowels, robust categorization was still invisible in the results of the ABX task. These empirical and conceptual considerations undermine the traditional assumption that robust categorization of a new phone contrast is a prerequisite for the establishment of corresponding lexical constrasts, but they are fully consistent with DMAP.

VIII. Conclusion

DMAP is motivated by strict conceptual necessity and finds support in our empirical results. A phonological account of spurious homophony due to merger finds support in the fact that lexical contrast and enhanced category distinctions do not go hand in hand. The acquisitional asymmetry between high and mid vowels is unexpected on the basis of acoustic or perceived differences and does not appear to merely reflect the familiarity of English-French learners with the test words since spurious homophony cuts across familiarity rates. It receives, however, a phonological account that underscores either contextual effects in view of the L1 phonological system or economy constraints on feature (re)assembly. Disentangling the role of these two factors in acquisition would involve comparing rounded vowel contrasts followed by rhotic and non-rhotic codas. These particular hypotheses will need to be verified with other stimuli and other learner groups.

DMAP provides a learning mechanism for the development of L2 phonological systems, which is in fact compatible with what is known about grammatical processing in general as well as relations between grammar and parsing. The phonological grammar provides a licensing mechanism for the representations extracted from the raw percepts. DMAP also characterizes aspects of the learning triggers for new category formation, with phonologically induced category formation and boundary shifts that would be long lasting, rather than short-lived, but involve the smallest change compatible with the new phonological state.

Crucially, DMAP challenges the assumption that category distinction precedes the development of an L2 phonemic inventory and lexical development, preserving however insights of the Perceptual Assimilation Model (Best and Tyler 2007) and Native Language Magnet Model (Kuhl and Iverson, 1995) as key components of DMAP for the development of category definitions. Many other issues arise, which can be experimentally tested.

References

Altenberg, E. P. and Vago, R. M. 1983: Theoretical implications of an error analysis of second language phonology production. Language Learning 33, 427-447.

Archibald, J. 1993: The learnability of English metrical parameters by adult Spanish speakers. IRAL 31, 129-142.

Archibald, J. 1998: Second language phonology, phonetics, and typology. Studies in Second Language Acquisition 20, 189-211.

Archangeli, Diana (1988). Aspects of Underspecification Theory. *Phonology*, 5, 183-207. Archibald, John (2004). Interfaces in the prosodic hierarchy: New structures and the phonological parser. *International Journal of Bilingualism* 8 (1), 29-50.

Avery, Peter, and Rice, Keren (1989). Segment structure and coronal underspecification. *Phonology*, 6, 179-2000.

Barlow, A. R., Pisoni, D. B., Akahane-Yamada, R. and Tohkura, Y. 1997: Training Japanese listeners to identify English /r/ and /l/, IV: some effects of perceptual learning on speech production. The Journal of the Acoustic Society of America 101, 2299-2310.

Best, C. T. 1995: A direct realist view of cross-language speech perception. In Strange, W., editor, Speech perception and linguistic experience: issues in cross-language research, Timonium, MD: York Press, 171-204.

Best, C. T., McRoberts, G. W, and Goodell, E. 2001: Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. Journal of the Acoustical Society of America 109, 775-794.

Best, C. T., McRoberts, G. W., and Sithole, N. M. 1988: Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. The Quarterly Journal of Experimental Psychology, A: Human Experimental Psychology 14, 345-360.

Best, C. T. and Tyler, M. D. 2007: Nonnative and second-language speech perception: commonalities and complementarities. In Bohn, O-S. and Munro, M. J., editors, Language experience in second language speech learning: in honor of James Emil Flege, Philadelphia: John Benjamins, 13-34.

Broselow, E. 1987: Non-obvious transfer: on predicting epenthesis errors. In Ioup, G. and Weinberger, S., editors, Interlanguage phonology, Cambridge, MA: Newbury House, 292-304.

Broselow, E., Chen, S.-I., and Wang, C. 1998: The emergence of the unmarked in second language phonology. Studies in Second Language Acquisition 20, 261-280.

Brown, C. 1998: The role of the L1 grammar in the L2 acquisition of segmental structure. Second Language Research 14, 136–193.

Brown, C. 2000: The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.), Second language acquisition and linguistic theory (pp. 4-63). Malden, Oxford: Blackwell Publishers.

Carlisle, R. 1998: The acquisition of consonants in a markedness relationship: A longuitudinal study. Studies in Second Language Acquisition 20, 245-260.

Clements, G. N. 1985: The geometry of phonological features. Phonology Yearbook 2, 225-252.

Clements, G. N. 2001. Representational economy in constraint-based phonology. In T. Alan Hall (Ed.), Distinctive feature theory (pp. 71–146). Berlin: Mouton de Gruyter.

Clements, G. N. 2003a. Feature economy in sound systems. Phonology 20, 287–333.

Clements, G. N. 2009. The role of features in phonological inventories. In Eric Raimy and Charles Cairns (Eds.), Contemporary views on architecture and representations in phonological theory, 19–68. Cambridge, Mass.: MIT Press.

Clements , G.N. and Hume, E. 1995. The Internal Organization of Speech Sounds. In J. Goldsmith (Ed.), Handbook of Phonological Theory (pp. 245-306). Oxford: Blackwell..

Cutler, A., Weber, A. and Otake, T. 2006: Asymmetric mapping from phonetic to lexical representations in second-language listening. Journal of Phonetics 34, 269–84.

Dehaene-Lambertz, G. 1997: Electrophysiological correlates of categorical phoneme perception in adults. NeuroReport 8, 919-924.

Dresher, B. E. (1999). Charting the learning path: Cues to parameter setting. *Linguistic Inquiry*, 30, 27 – 67.

Dresher, B. E. and Kaye, J. D. (1990). A computational learning model for metrical phonology. *Cognition* 34,137–195

Dresher, B. Elan. 2010. Feature Hierarchies and Contrast in Phonology. In Greville G. Corbett and Anna Kibort, eds., *Features: Perspectives on a Key Notion in Linguistics*, 37-63. Oxford: Oxford University Press.

Dufour, S., Nguyen, N., and Frauenfelder, U. H. 2007 : The perception of phonemic contrasts in a non-native dialect. The Journal of the Acoustical Society of America 121, EL131-EL136.

Dupoux, E., Pallier, C., Kakehi, K., and Mehler, J. 2001: New evidence for prelexical phonological processing in word recognition. Language and Cognitive Processes 16, 491-505.

Eckman, F. R. 1977: Markedness and the contrastive analysis hypothesis. Language Learning 27, 315–330.

Eckman, F. R. 1987: Markedness and the Contrastive Analysis Hypothesis. In Ioup, G. and Weinberger, S. H., editors, Interlanguage Phonology, Cambridge, MA: Newbury House, 55-69.

Eimas, P. D. 1974: Auditory and linguistic processing of cues for place of articulation by infants. Perception and Psychophysics 16, 513-521.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., and Vigorito, J. 1971: Speech perception in infants. Science 171, 303-306.

Eisner, F., and McQueen, J. M. 2005: The specificity of perceptual learning in speech processing. Perception and Psychophysics 67, 224-238.

Escudero, P., Hayes-Harb, R., and Mitterer, H. 2008: Novel second-language words and asymmetric lexical access. Journal of Phonetics 36, 345-360.

Evans, B. and Iverson, P. 2002: Vowel normalization for accent: an investigation of best exemplar locations in northern and southern British English sentences. Speech, Hearing and Language: work in progress. Department of Phonetics and Linguistics, University College London 14, 18-38.

Fant, G. 1969. Distinctive features and phonetic dimensions. STL-QPSR, 10, 01-18.

Flege, J. E. 1987: The production of ‘‘new’’ and ‘‘similar’’ phones in a foreign language: evidence for the effect of equivalence classification. Journal of Phonetics 15, 47–65.

Flege, J. E. 1995: Second language speech learning: theory, findings, and problems. In Strange, W., editor, Speech perception and linguistic experience: issues in cross-language research, Timonium, MD: York Press, 233-277.

Flege, J. E. and Hillenbrand, J. 1984: Limits on phonetic accuracy in foreign language speech production. Journal of the Acoustical Society of America 76, 708–721.

Francis, A. L., Baldwin, K., and Nusbaum, H. C. 2002: Effects of training on attention to acoustic cues. Perception & Psychophysics 62, 1668-1680.

Francis, A. L., and Nusbaum, H. C. 2002: Selective attention and the acquisition of new phonetic categories. Journal of Experimental Psychology: Human Perception and Performance 28, 349-366.

Fujisaki, H., and Kawashima, T. 1971: A model of the mechanisms for speech perception: Quantitative analyses of categorical effects in discrimination. Annual Report of the Engineering Research Institute, Faculty of Engineering, University of Tokyo, pp. 59-68.

Goad, H., and White, L. 2006: Ultimate attainment in interlanguage grammars: A prosodic approach. Second Language Research 22, 243-268.

Goad, H., and White, L. 2008. Prosodic structure and the representation of L2 functional morphology: A nativist approach. Lingua 118, 577-594.
Goad, H., White, L. and Steele, J. 2003: Missing inflection in L2 acquisition: Defective syntax or L1-constrained prosodic representations? Canadian Journal of Linguistics 48, 243-263.
Gottfried, T. L. 1984: Effects of consonant context on the perception of French vowels. Journal of Phonetics 12, 91-114.
Halle, M. 1992: Phonological features. In W. Bright (ed.) International encyclopedia of linguistics. New York: Oxford University Press. 207-212.
Halicki, S. D. 2010: Learner knowledge of target phonotactics: Judgements of French word transformations. LINCOM Europa
Hancin-Bhatt, B. 1994: Segment transfer: a consequence of a dynamic system. Second Language Research 10, 241-269.
Hayes-Harb, R., and Masuda, K. 2008: Development of the ability to lexically encode novel second language phonemic contrasts. Second Language Research 24, 5-33.
Hillenbrand, J. M, Clark, M. J., and Nearey, T. M. 2001: Effects of consonant environment on vowel formant patterns. Journal of the Acoustical Society of America, 109, 748-763.
Højen, A., and Flege, J. E. 2006: Early learners' discrimination of second-language vowels. Journal of the Acoustical Society of America, 119, 3072-3084.
Hume, E. and Bromberg, I. 2005: Predicting epenthesis: an information-theoretical account. 7th Annual Meeting of the French Network Phonology. Aix-en-Provence, June 2005.
Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. and Siebert, C. 2003: A perceptual interference account of acquisition difficulties for non-native phonemes. Cognition 87, B47-B57.
Keyser, S. J., and Stevens, K. N. 1994: Feature Geometry and the Vocal Tract. Phonology, 11(2), 207-236.
Kuhl, P. K. and Iverson, P. 1995: Linguistic experience and the “perceptual magnet effect”. In Strange W., editor, Speech perception and linguistic experience: issues in cross-language research, Timonium, MD: York Press, 121-154.
Kuhl, P., Williams, K., Lacerda, F., Stevens, K., and Lindblom, B. 1992: Linguistic experience alters phonetic perception in infants by 6 months of age. Science 255, 606-608.
Lardiere, D. 2009: Some thoughts on the contrastive analysis of features in second language acquisition. Second Language Research 25: 173-227.
Levy, E. S. 2009a: Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. Journal of the Acoustical Society of America, 125, 1138-1152.
Levy, E. S. 2009b: On the assimilation-discrimination relationship in American English adults' French vowel learning. Journal of the Acoustical Society of America, 126, 2670–2682.

Levy, E. S., and Strange, W. 2008: Perception of French vowels by American English adults with and without French language experience. Journal of Phonetics 36, 141-157.

Liberman, A. M., Harris, K., Hoffman, H. S., and Griffith, B. 1957: The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology 54, 358-368.

Lively, S. E., Logan, J. S., Pisoni, D. B. 1993: Training Japanese listeners to identify English /r/ and /l/, II: the role of phonetic environment and talker variability in learning new perceptual categories. The Journal of the Acoustic Society of America 94, 1242-1255.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y. and Yamada, T. 1994: Training Japanese listeners to identify English /r/ and /l/, III: long-term retention of new phonetic categories. The Journal of the Acoustic Society of America 96, 2076-2087.

Logan, J. S., Lively, S. E. and Pisoni, D. B. 1991: Training Japanese listeners to identify English /r/ and /l/, I: a first report. The Journal of the Acoustic Society of America 89, 874.

Logan, J. S., Lively, S. E. and Pisoni, D. B. 1993: Training listeners to perceive novel phonetic categories: how do we know what is learned? The Journal of the Acoustic Society of America 94, 1148.

Magnan, S. 1986: Assessing speaking proficiency in the undergraduate curriculum: data from French. Foreign Language Annals 5, 429-438.

Maye, J. 2000: The acquisition of speech sound categories on the basis of distributional information, unpublished doctoral dissertation, University of Arizona.

Maye, J., Werker, J. F. and Gerken, L. 2002: Infant sensitivity to distributional information can affect phonetic discrimination. Cognition 82, B101-B111.

Mielke, J., (2008) The emergence of distinctive features. Oxford University Press, USA.

McCarthy, J. J. 1988: Feature geometry and dependency: a review. Phonetica 45, 84-108.

McQueen, J. M., Cutler, A., and Norris, D. 2003: Flow of information in the spoken word recognition system. Speech Communication 41, 257-270.

McQueen, J. M., Cutler, A., and Norris, D. 2006: Phonological abstraction in the mental lexicon. Cognitive Science 30, 1113-1126.

New B., Pallier C., Ferrand L., Matos R. 2001 : Une base de données lexicales du français contemporain sur internet: LEXIQUE, L'Année Psychologique, 101, 447-462. http://www.lexique.org

New, B., Pallier, C., Brysbaert, M., Ferrand, L. 2004: Lexique 2 : A New French Lexical Database. Behavior Research Methods, Instruments, & Computers, 36 (3), 516-524.

Ohala, J. 1995: Phonetic explanations for sound patterns: implications for grammars of competence. ICPhS (2), 52-59.

Ota, M., Hartsuiker, R. J., and Haywood, S. L. 2009: The KEY to the ROCK: near-homophony in nonnative visual word recognition. Cognition, 111, 263-269.

Pallier, C., Bosch, L. and Sebastián-Gallés, N. 1997: A limit on behavioral plasticity in vowel acquisition. Cognition 64, B9-B17.

Pallier, C., Colomé, A. and Sebastián-Gallés, N. 2001: The influence of native-language phonology on lexical access: exemplar-based vs. abstract lexical entries. Psychological Sciences 12, 445-449.

Pater, J. 2003: The perceptual acquisition of Thai phonology by English speakers: task and stimulus effects. Second Language Research 19, 209-223.

Pisoni, D. B. 1973: Auditory and phonetic memory codes in the discrimination of consonants and vowels. Perception & Psychophysics 13, 253-260.

Polka, L. 1995: Linguistic influences in adult perception of non-native vowel contrasts. Journal of the Acoustical Society of America 97, 1286-1296.

Rialland, A. 1986: Schwa et syllabes en Français. In Wetzels, L. and Sezer E. (Eds.) *Studies in Compensatory Lengthening*, (pp. xxxx). Foris Publications.

Schwartz, B. D. and Sprouse, R. A. 1994: Word order and nominative case in non-native language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In Hoekstra, T. and Schwartz, B. D., editors, Language acquisition studies in generative grammar: papers in honor of Kenneth Wexler from the 1991 GLOW workshops, Philadelphia, PA: John Benjamins 317-368.

Schwartz, B. D. and Sprouse, R. A. 1996: L2 cognitive states and the Full Transfer/Full Access model. Second Language Research 12, 40-72.

Sebastián-Gallés, N., Echeverria, S. and Bosch, L. 2005: The influence of initial exposure on lexical representation: comparing early and simultaneous bilinguals. Journal of Memory and Language 52, 240-255.

Stevens, K. N., Libermann, A. M., Studdert-Kennedy, M. and Öhman, S. E. G. 1969: Crosslanguage study of vowel perception. Language & Speech 12, 1-23.

Strange, W., Weber, A., Levy, E., S., Shafiro, V., Hisagi, M., and Nishi, K. 2007: Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. Journal of the Acoustical Society of America, 122(2), 1111-1129.

Weber, A., and Cutler, A. 2004: Lexical competition in non-native spoken-word recognition. Journal of Memory and Language 50, 1-25.

Werker, J. F., and Curtin, S. 2005: PRIMIR: a developmental framework of infant speech processing. Language learning and development 1, 197-234.

White, L., Melhorn, J. F. and Mattys, S. L. 2009: Segmentation by lexical subtraction in Hungarian speakers of second-language English. The Quarterly Journal of Experimental Psychology.

Young-Scholten, M. 2004: Prosodic constraints on allophonic distribution in adult L2 acquisition. International Journal of Bilingualism 8, 67-77.

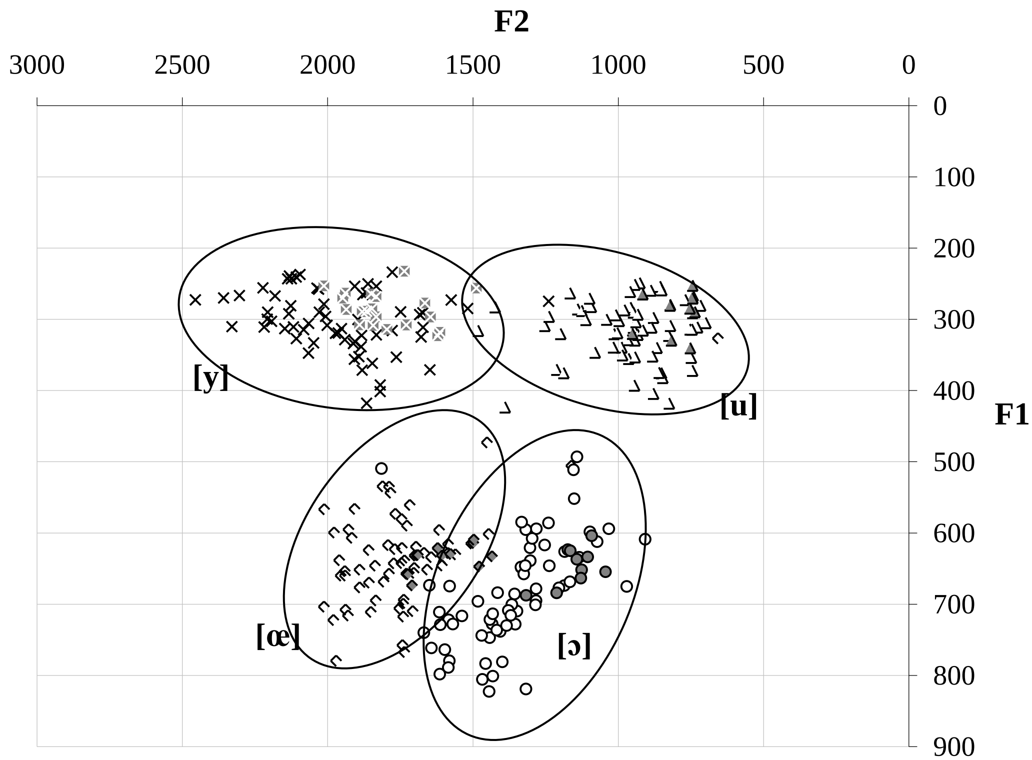

Figure 1. F1 and F2 frequencies measured at the midpoint for the critical vowels in each of the /y/-/u/ and /œ/-/ɔ/ minimal pairs for ABX (open symbols) and lexical decision (filled symbols).

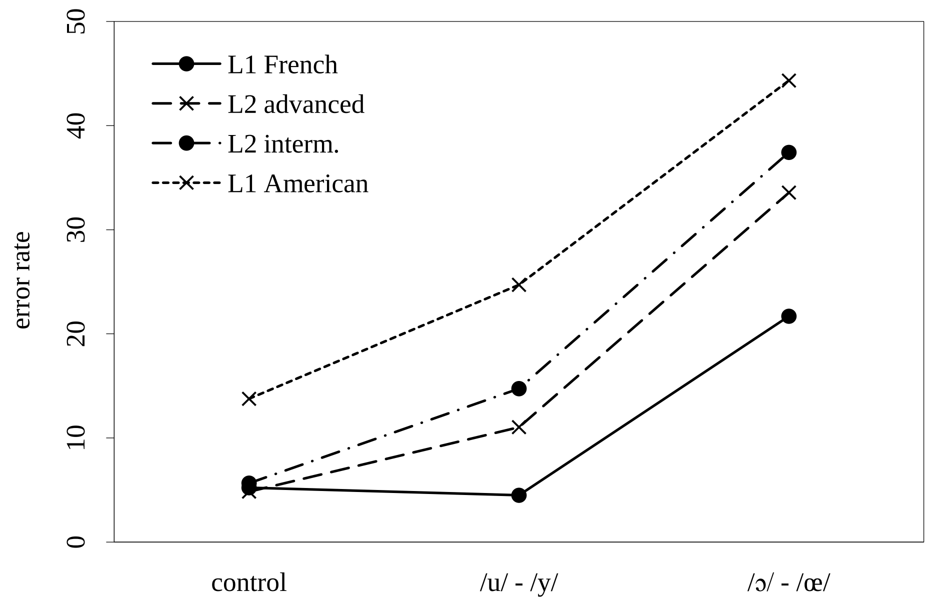

Figure 2. Error rate by condition (control, high /y/-/u/, mid /œ/-/ɔ/) and group (monolingual English native speakers, intermediate L2 learners, advanced L2 learners and native French speakers)

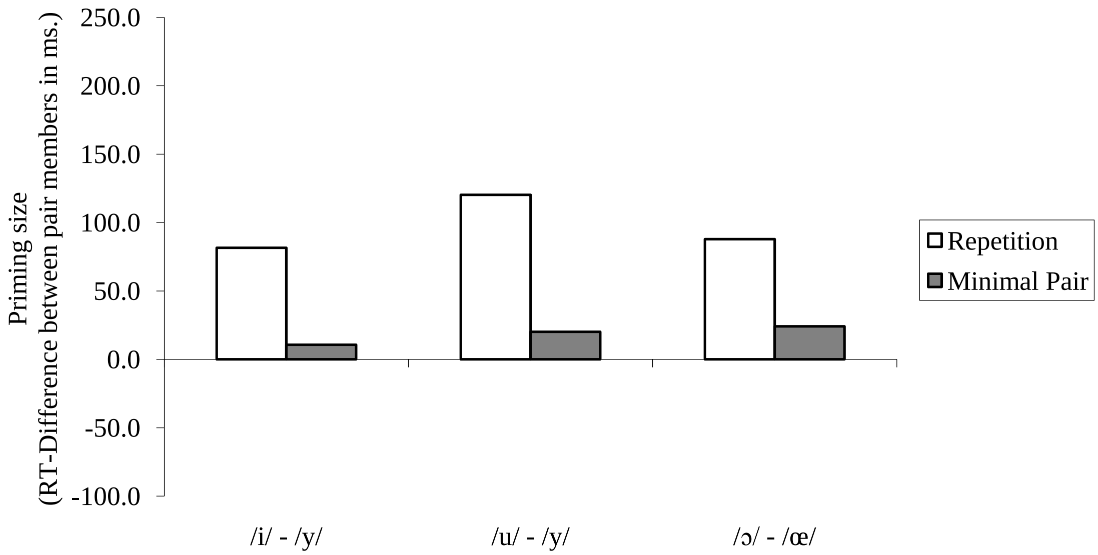

Figure 3. Priming obtained in each condition and contrast in the native speaker group.

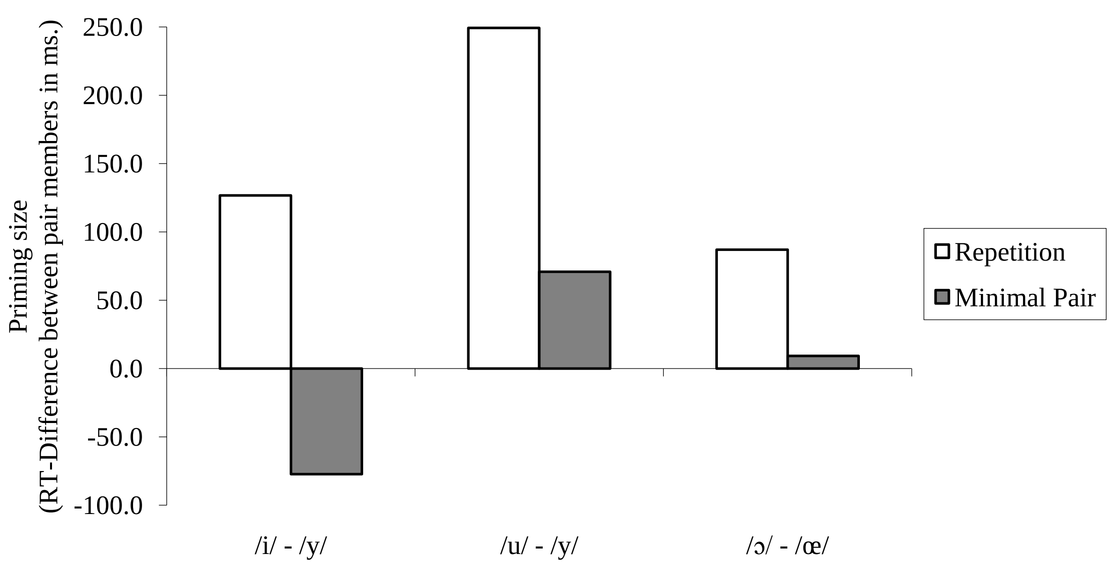

Figure 4. Priming obtained in each condition and contrast in the advanced learner group.

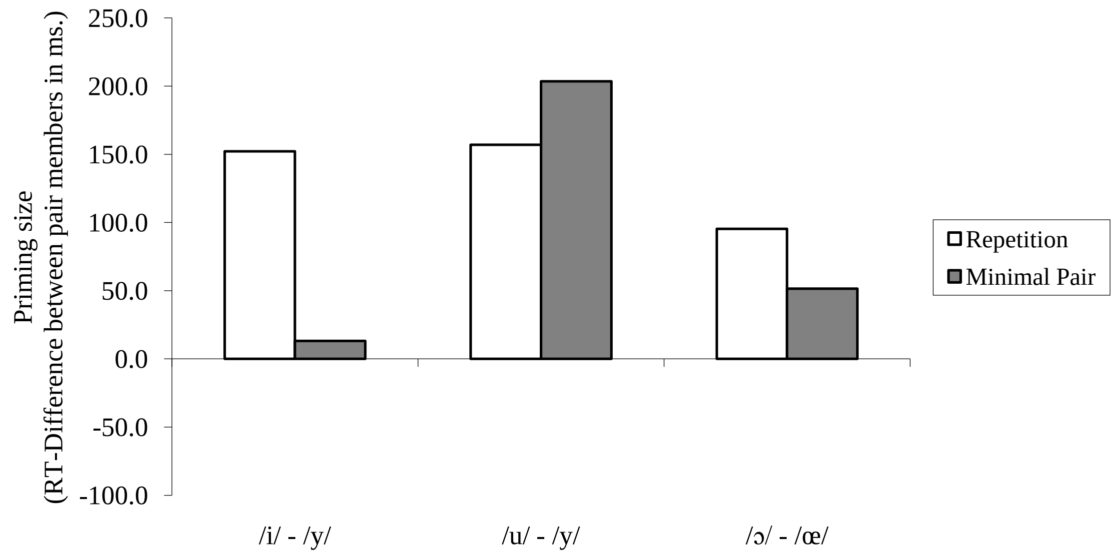

Figure 5. Priming obtained in each condition and contrast in the intermediate learner group.

Table 1: overview of frequency distribution for the contrasts used in the minimal pairs (in occurrences per million)

	written (freqlivres)	**spoken (freqfilms2)**	**number of homophonic words (nbhomoph)**
/y/-/u/	73.77	41.41	7.0
/œ/-/ɔ/	71.58	40.71	4.9
/i/-/y/	27.96	31.40	6.4

Table 2: duration measurements for all stimuli (in ms.): The items that have a lengthening coda are measured separately, for both /œ/ and /ɔ/ (L-only)

	LD	ABX	*p*-value	
/œ/	223	170	0.02	*
/œ/ (L-only)	(223)	256	0.19	ns
/ɔ/	224	167	0.02	*
/ɔ/ (L-only)	(224)	252	0.31	ns
/u/	153	165	0.40	ns
/y/	148	170	0.07	ns

Table 3: error rates as a function of consonantal context in each group

High vowels (u-y)	**Labial**	**Coronal**	***p*-value**
French natives	4.6	4.3	>.1
Advanced	9.7	12.3	$t(18) = 2.8$ $p < .07$
Intermediate	12.0	17.4	$t(18) = 2.8$ $p < .005$
English monolinguals	21.6	27.8	$t(12) = 2.7$ $p < .01$
Mid vowels (ɔ-œ)	**Labial**	**Coronal**	***p*-value**
French natives	21.1	22.2	>.1
Advanced	31.7	35.3	>.1
Intermediate	34.2	40.6	$t(18) = 2.2$ $p < .02$
English monolinguals	39.1	49.5	$t(12) = 2.9$ $p < .006$

Table 4: Error rate (in %) according to lengthening context in ABX; Mid = mid vowels, High = high vowels

	Mid-L	Mid-NL	High-L	High-NL
L1-French	18.1	23.3	5.0	4.3
Advanced	31.6	34.5	9.7	11.6
Intermediate	34.2	35.2	15.5	14.1
L1-American	42.6	45.1	25.6	24.3

Table 5: Reaction times (SE) and priming size by group and contrast in each condition

		Intermediates			Advanced			Native speakers		
		Mean RT	(SE)	Priming	Mean RT	(SE)	Priming	Mean RT	(SE)	Priming
i-y	rep	1170.8	38.4		1009.7	31.7		816.8	24.7	
	rep	1018.7	48.9	152.2	883.1	36.3	126.7	735.3	42.4	81.5
	mp	1103.3	45.3		996.8	51.1		832.3	25.3	
	mp	1090.1	50.8	13.2	1074.1	46.1	-77.3	821.6	41.4	10.6
y-u	rep	1129.6	38.5		1154.8	41.4		872.1	31.0	
	rep	972.7	34.0	156.9	905.6	28.9	249.2	751.9	30.5	120.2
	mp	1240.8	52.8		1128.6	41.8		873.5	29.5	
	mp	1037.3	30.7	203.6	1057.8	34.9	70.8	853.4	29.7	20.1
œ-ɔ	rep	1017.7	41.3		1021.7	31.1		825.6	34.3	
	rep	922.4	35.4	95.4	934.7	40.5	87.0	737.8	41.8	87.9
	mp	1071.2	42.4		1046.6	41.3		863.4	44.0	
	mp	1019.6	37.4	51.5	1037.4	55.7	9.3	839.3	38.9	24.1

Appendix

Test words used in Lexical Decision for the contrasts /y/-/i/, /y/-/u/, /œ/-/ɔ/

/y/			/i/		
bûche	[byʃ]	‘log’	biche	[biʃ]	‘deer’
cru	[kʁy]	‘raw/believed’	cri	[kʁi]	‘scream’
dune	[dyn]	‘dune’	dine	[din]	‘dine’
ruche	[ʁyʃ]	‘beehive’	riche	[ʁiʃ]	‘rich’
vue	[vy]	‘seen’	vie	[vi]	‘life’

/y/			/u/		
bulle	[byl]	‘bubble’	boule	[bul]	‘ball’
puce	[pys]	‘flea’	pouce	[pus]	‘thumb’
rue	[ʁy]	‘street’	roue	[ʁu]	‘wheel’
bu	[by]	‘drunk’	bout	[bu]	‘piece’
sur	[syʁ]	‘sure/on’	sourd	[suʁ]	‘deaf’

/œ/			/ɔ/		
beurre	[bœʁ]	‘butter’	bord	[bɔʁ]	‘border’
coeur	[kœʁ]	‘heart’	corps	[kɔʁ]	‘body’
peur	[pœʁ]	‘fear’	port/porc	[pɔʁ]	‘harbor/pork’
seul	[sœl]	‘alone’	sol/sole	[sɔl]	‘floor/sole’
leur	[lœʁ]	‘their’	lors	[lɔʁ]	‘then’

Test nonwords used in Lexical Decision for the contrasts /y/-/i/, /y/-/u/, /œ/-/ɔ/

/y/		/i/	
duge	[dyʒ]	dige	[diʒ]
nur	[nyʁ]	nir	[niʁ]
plune	[plyn]	pline	[plin]
cruffe	[kʁyf]	criffe	[kʁif]
vurte	[vyʁt]	virte	[viʁt]

/y/		/u/	
puche	[pyʃ]	pouche	[puʃ]
chupe	[ʃyp]	choupe	[ʃup]
bluche	[blyʃ]	blouche	[bluʃ]
frue	[fʁy]	froue	[fʁu]
bruse	[bʁyz]	brouse	[bʁuz]

/oe/		/ɔ/	
bleuve	[blœv]	blove	[blɔv]
jeur	[ʒœʁ]	jore	[ʒɔʁ]
breuf	[bʁœf]	broffe	[bʁɔf]
creulle	[kʁœl]	crolle	[kʁɔl]
deusse	[dœs]	dosse	[dɔs]

ABX Stimuli (pairs of French non-words)

	labial		coronal	
œ-ɔ	bœb	bɔb	lœd	lɔd
	bœp	bɔp	lœl	lɔl
	fœb	fɔb	lœt	lɔt
	fœp	fɔp	nœz	nɔz
	fœv	fɔv	sœz	sɔz
	mœp	mɔp	tœd	tɔd
	vœb	vɔb	tœz	tɔz
	vœp	vɔp	zœz	zɔz
u-y	bub	byb	lud	lyd
	bup	byp	lul	lyl
	fub	fyb	lus	lys
	fup	fyp	nuz	nyz
	fuv	fyv	suz	syz
	mup	myp	tud	tyd
	vub	vyb	tuz	tyz
	vup	vyp	zuz	zyz
Control vowel /i/-/ɛ/	bif	bɛf	dil	dɛl
	pim	pɛm	ziz	zɛz
	bip	bɛp	niz	nɛz
	mip	mɛp	tid	tɛd
Control consonant (across both contexts)		bup	tud	
		myb	lys	
		fub	nut	
		vyb	syn	
		pum	nus	
		vyp	lyd	
		mup	zun	
		fyp	lyl	

[1] This could not be avoided due to the restrictions of using the same CVC combinations for both vowel pairs. This and a few other items were close to English words; however, the French phonetics of the stimuli (e.g. [ɔ] is not a native vowel for our participants) is likely to prevent strong activation of English vocabulary.