Phonological knowledge in compensation for native and non-native assimilation

I. Darcy[1,2]; F. Ramus[1,3]; A. Christophe[1]; K. Kinzler[4]; E. Dupoux[1]

[1]Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole des Hautes Etudes en Sciences Sociales, Centre National de la Recherche Scientifique, Ecole Normale Supérieure, Paris, France

[2] Institut für Linguistik, Universität Potsdam, Germany

[3] Institute of Cognitive Neuroscience, University College London, UK

[4] Laboratory for Developmental Studies, Harvard University, USA

Short title : native and non-native compensation

keywords : voice and place assimilation, phonology, speech perception, lexical recognition

Corresponding Author:
Isabelle Darcy
Universität Potsdam
Institut für Linguistik
Postfach 601553
D-14415 Potsdam
Germany
tel: +49 331 977 2401
fax: +49 331 977 2087
e-mail: darcy@ling.uni-potsdam.de

Abstract

We investigated whether compensation for phonological assimilation in the first language depends or not on language specific knowledge of phonological processes. To this end, we tested two different assimilation rules, one that exists in English and involves place of articulation, and another that exists in French and involves voicing. Both contrasts were tested on speakers of French and American English. In two experiments using a word detection task, we observed that participants showed a significantly higher degree of compensation for phonological changes that correspond to rules existing in their language than to rules that do not exist in their language (even though they are phonologically possible since they exist in another language). Thus, French participants compensated more for voicing than place assimilation, while American English participants compensated more for place than voicing assimilation. In both experiments, we also found that the non-native rule induced a very small but significant compensation effect, suggesting that both a language-specific and a language-independent mechanism are at play. Control experiments ensured that changes in stimuli were perceived clearly in isolation, differential compensation then being due to the phonological context of change, rather than to specific phonetic cues. The results are discussed in light of current models of lexical access and phonological processing.

Introduction

Understanding how words are recognized in continuous speech presents a particular challenge because the acoustic and phonetic shape of a word may be severely distorted in continuous speech compared to when that word is spoken in isolation. Words in sentences can be up to twice as short as words spoken in citation form. This higher speaking rate results in a number of acoustic changes due to co-articulation between the segments within and between the words (Church, 1987; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Trubetzkoy, 1958). Even more dramatic, some language-specific phonological rules substitute, insert or delete entire segments as a function of speaking rate or phonological context (see Table 1). Such changes can potentially disrupt lexical recognition, since they can neutralize existing contrasts between phonemes, and hence contrasts between lexical items. In English, for example, place assimilation affects coronal stops, which take on the place of articulation of the following stop in connected speech (Barry, 1992; Ellis & Hardcastle, 2002; Nolan, 1992). Hence the compound *football* may be realized as *foo[p]ball*. In French, voicing (glottal) assimilation voices obstruents before voiced obstruents, and devoices them before unvoiced obstruents (Dell, 1995; Féry, 2003, Wetzels & Mascaró, 2001, Snoeren, Hallé & Segui, 2005). So, the same word *football* tends to be realized as *foo[d]ball*. Such rules are common across the world's languages and tend to be productive, applying to novel items. Moreover, when several rules coexist in a language, they can be chained to one another, resulting in large changes in surface word forms. For instance in French, the rules of nasal-obstruent simplification or word-final liquid deletion (Casagrande, 1984; Dell, 1995; Féry, 2003) can be chained with regressive glottal assimilation: the sequence *table carrée* [tabl+kaʁe] 'square table' can thus become [tapkaʁe] in casual speech.

Even though there is considerable debate in the phonetic literature as to whether the phonetic change is complete or leaves traces of the original segment (Ellis & Hardcastle, 2002; Féry, 2003; Nolan, 1992), it remains true that these rules substantially affect the phonetic shape of words. This in turn may render the identification of lexical entries problematic. The surprising fact is that these phonological changes seem to matter very little in everyday continuous speech recognition. In fact, most people are not even aware of the existence of these phonological changes. This calls for an explanation. What are the mechanisms responsible for robust lexical access despite near neutralizing changes induced by phonological rules?

We review three classes of mechanisms that have been proposed in the literature. We call them *lexical compensation*, *phonetic compensation* and *language-specific phonological inference*. We don't assume that models presented within a section are interchangeable. We discuss together proposals that make similar predictions regarding three crucial features of compensation, which we will explain in the course of the review. The purpose of this paper is not primarily to distinguish between processing architectures or modeling details (which would require much more experiments), but rather to understand more in depth some aspects of compensation, given contradictory evidence in the literature.

Table 1. Examples of phonological rules which change the shape of words according to phonological context.

Type of rule	Source	Rule	Language	Example
Regressive devoicing	Dell, 1995; Féry, 2003	$Obs_{[+vd]} \rightarrow Obs_{[-vd]}$ / __ (#)$Obs_{[-vd]}$	French	*robe sale* ro/**b**#s/ale 'dirty dress' → [ʀɔ**p**sal]
Progressive devoicing	Wetzels & Mascaro, 2001	$Obs_{[+vd]} \rightarrow Obs_{[-vd]}$ / $Obs_{[-vd]}$ (+) __	Dutch	/v/allen 'to fall' o/p+**v**/allen 'to strike' → o[pf]allen
Regressive place	Wells, 1982, p. 55	$C_{[cor]} \rightarrow C_{[vel]}$ / _ (#)$C_{[vel]}$	English	*good girl* goo/**d**#g/irl → [gʊggɜːl]
Progressive place	Wiese, 1996	$C_{[cor]} \rightarrow C_{[lab]}$ / $C_{[lab]}$ __	German	*geben* 'to give' /gebən/ → [gebm̩] *halten* 'to hold' /haltən/ → [haltn̩]
[back] Harmony	Roca & Johnson, 1999, p.154	$V \rightarrow V_{[+ back]}$ / $V_{[+ back]}$ C_0 + C_0 __ C_0#	Turkish	[ip+in] 'rope' (Gen.sg.) ; [sap+ɯn] 'stalk' (Gen.sg.)
Liquid deletion	Dell, 1995	Liq → ∅ / Obs_ # Obs	French	*table jaune* 'yellow table' /tab**l**#ʒon/ → [tabʒon]
Nasal-obstruent simplification	Féry, 2003	ṼObs → ṼN / _(#)N	French	*Langue maternelle* 'native language' lan/g#m/aternelle → lan[ŋm]aternelle
Consonant insertion	Wells, 1982, p. 58	∅ → r / V(#)__ (#)V (-high)	English	*Magnoli**a** **in** the garden* → magnoli[əɹin]… or *sofa is* /sofa#iz/ → [sofəɹis]

Lexical Compensation

The first class of compensation mechanisms uses lexical knowledge. Since we know the words in our language, we can match the incoming signal with our stored list and pick the closest and/or most likely candidate available. This strategy essentially treats phonetic variation as random noise, and uses lexical and higher-order context to recover the signal from that noise. It is actually put to use in several speech recognition systems, and their mere existence attests the feasibility of such a mechanism. There is some evidence in psycholinguistics that lexical access incorporates robust mechanisms that resist input degradation. For instance, in running speech, lexical recognition resists mispronunciations; participants might even have a difficult time to detect mispronunciations in fluent speech (Marslen-Wilson & Welsh, 1978), and 'hallucinate' phonemes replaced by noise on the basis of lexical (and phonetic) proximity (Samuel, 1981, 1996, 2001). Recent models of lexical recognition have implemented such robustness by relying on multiple activation of lexical candidates and competition between them (see the *Cohort* model, Marslen-Wilson & Welsh, 1978; the *TRACE* model, McClelland & Elman, 1986; and *Shortlist*, Norris, 1994). This insures that whenever a degraded input is presented, several lexical candidates will be activated. Lexical competition, plus potentially higher-order expectations, ensures that the most plausible candidate is finally selected (Gow & Gordon, 1995).

Although mechanisms like phoneme restoration may account for part of phonological compensation effects, they fail to distinguish between one-feature mispronunciations (which are often noticed) and one-feature assimilations (which are hardly ever noticed). Lahiri and Marslen-Wilson (1991, 1992) therefore developed a model of compensation based on underspecification theory (Archangeli, 1988; Kiparsky, 1985; Pulleyblank, 1988), which explicitly implements regular phonological variation within lexical representations: They assume featurally "underspecified" lexical representations for words (FUL, Featurally Underspecified Lexicon, see Lahiri & Reetz, 2002), for precisely those features that display regular variation. For instance, in English coronal stops would be unspecified for place, whereas labial or velar stops would be specified for place. Words containing coronal stops would thus have a gap in their featural specification; as a consequence, a deviant phonetic input could be mapped onto an unspecified segmental slot. Therefore, even if the sensory input differed in one position by one feature, its representation could nevertheless activate the appropriate lexical entry (see also Marslen-Wilson, Nix, & Gaskell, 1995). This theory predicts an asymmetry in the recognition of lexical items depending on whether or not they contain unspecified segments. Using cross-modal priming, Lahiri and Reetz found that the deviant nonword stimulus **Bah[m]* triggered as much priming for the target *Zug* ('train' semantically related to *Bahn*) as the unchanged word *Bahn* 'railway' (where the coronal /n/ is assumed to be unspecified for place). In contrast, and consistent with their prediction, the deviant stimulus **Lär[n]* did not prime the target *Krach* 'bang', whereas the unchanged word *Lär[m]* 'noise' did (/m/ being specified as labial, only labials could map onto this slot). Note however that this result was not replicated by Gow (2001) who found equal priming for two similar conditions in English. Although the underspecification model cannot be fully equated with other models of lexical compensation, the predictions of all these models are similar.

Lexical compensation mechanisms have two crucial features. First, they rely on stored lexical items, and hence only work for restoring the phonological shape of actual words – not nonwords. Second, in their rudimentary form, they are insensitive to phonological context: the best-matching lexical item is selected based on the local phonetic cues and optionally the semantic and/or syntactic context. Crucially for the present experiments, the activation and selection of the most appropriate lexical item does not take into account the phonological context in which the changes occur, and whether these changes are systematic in the language or not.

Regarding the first feature (compensation for nonwords), most studies have used real words to assess compensation for assimilation. Using phoneme detection though, Gaskell and Marslen-Wilson (1998) found results with nonwords that were parallel to those of real words, although the amplitude of the effect was smaller. This effect on nonwords is impossible to account for with lexical compensation and suggests that compensation for assimilation is at least partly due to a non-lexical mechanism (see also Gaskell, Hare, & Marslen-Wilson, 1995; Mitterer & Blomert, 2003; Mitterer, Csépe, & Blomert, 2003).

Regarding the second feature (sensitivity to context), there is some robust evidence that compensation is sensitive to the segmental context in which the change occurs. For instance, Gaskell and Marslen-Wilson (1996) used cross-modal priming to examine compensation for place assimilation in English and observed more priming when the context was viable (leam#bacon → LEAN) than when it was unviable (leam#gammon → LEAN). These results were replicated and extended using other methods and assimilation processes by Coenen, Zwitserlood and Bölte (2001), as well as by Mitterer and colleagues (Mitterer & Blomert, 2003; Mitterer et al., 2003). This sensitivity to context is not predicted by the compensation model based on underspecification (see above), where **Bah[m]* is expected to be recognized as a token of *Bahn*, without any influence of the context. In sum, it seems that one crucial property of lexical compensation mechanisms, i.e. insensitivity to phonological context, does not hold for phonological compensation.

Instead, the use of context for compensation could originate in perceptual salience sensitivity which would be different across phonetic contexts, or it could also reflect the application of a kind of phonological knowledge, e.g. a familiarity with this particular type of modification (language-specific knowledge of the processes at work in a given language). The first possibility predicts that context effects and compensation are to be found also for processes that don’t exist in the language, as long as the appropriate context is given. The second one limits compensation phenomena and context effects to those processes that exist in a language. Contradictory results in the literature mirror a vivid debate as to whether compensation reflects language-specific knowledge or not. This third crucial feature is exactly the point of divergence between the two remaining classes of models.

Let us review first some of the evidence in favor of phonetic compensation, which is not dependent on language-specific processing, but rather takes place at a lower level of processing.

Phonetic compensation

This class of compensation mechanisms is based on acoustic/phonetic processes. The idea is to deal with compensation for phonological variation using those mechanisms that compensate for phonetic variation or coarticulation. Several decades of research in acoustic/phonetics have shown that acoustic cues relevant to a given segment are temporally spread out across adjacent positions (Bailey & Summerfield, 1980; Stevens, 1998). It has also been shown that the perceptual apparatus of listeners integrates multiple cues to the same feature (Best, Morrongiello, & Robson, 1981; Hodgson & Miller, 1996; Parker, Diehl, & Kluender, 1986; Repp, 1982; Sinnott & Saporita, 2000; Summerfield & Haggard, 1977; Treiman, 1999). These effects seem to hold across languages, and might even not be specific to humans, since compensation for coarticulation has been observed in birds (e.g. Lotto, Kluender, & Holt, 1997).

Gow (2001, 2002a, 2003) proposed a language independent processing mechanism called *Feature Cue Parsing* to handle both coarticulation and systematic phonological variation. In this mechanism, temporally distributed acoustic cues of feature values are grouped and integrated into segmentally aligned phonetic features (see also Fowler, 1996; Fowler & Brown, 2000). Gow's specific proposal is that feature parsing can account both for coarticulatory compensation and compensation for phonological assimilation, at least in the (frequent) cases where assimilation is not complete. Indeed, in most cases, the target phoneme contains phonetic traces or partial cues of the original unassimilated form (Ellis & Hardcastle, 2002; Nolan, 1992). The principle of feature parsing is the following: Complex segments that simultaneously encode two places of articulation are parsed onto two adjacent segmental positions, when the following context may attract one of the features. Attraction may take place when the following segment shares the same place of articulation as one of the two encoded in the preceding segment (Gow & Zoll, 2002, example (2) p.58). As a result, feature parsing may suffice to give an account of compensation for phonological rules, because the information used to parse the input is provided by the phonetic signal alone. For this same reason, this process is assumed to be language-independent. Supporting evidence is found in Gow (2001, exp. 1), where one existing process (place assimilation from coronals to labials, e.g. *green* becoming [grim]) was tested against a non-existing one (place assimilation from labials to coronals, e.g. *glum* becoming [glun]). No effect linked to experience with a given phonological assimilation process emerged (same priming effect in a lexical decision task, see also Gumnior et al., 2005, for a similar lack of context effect in German).

Note that although Feature Parsing may work when assimilation is incomplete, it does not provide an appropriate explanation when assimilation is complete: in this case, articulatory features are not spread across adjacent segments. Yet, several experiments have shown that compensation does occur with tokens that were deliberately produced with complete assimilation of the target phoneme (Coenen et al., 2001; Gaskell & Marslen-Wilson, 1996, 1998; Mitterer & Blomert, 2003). Further, Nolan (1992) and Ellis and Hardcastle (2002) demonstrated that a substantial proportion of spontaneous place assimilatory changes in English seem to be complete: that is, they left no detectable acoustic traces of the underlying phoneme. In addition, Feature Parsing would have trouble handling cases in which assimilation skips over 'transparent' consonants, like [m] in the Russian phrase /iz mtsenska/ [i**s** mtsenska] 'from Mcensk' (Hayes, 1984, Jakobson, 1956). Similarly, cases where listeners are confronted to elision, insertion or a combination of several processes would be hard to explain. Thus, although the Feature Parsing model could account for cases of partial assimilation, it does not seem to be powerful enough, or abstract enough, to deal with the full spectrum of phonological variation.

Language-specific phonological inference

A third class of mechanisms has been proposed to deal specifically with phonological sources of variation: phonological inference. This was first developed in Marslen-Wilson et al. (1995). Basically, phonological inference would be a language-specific mechanism that undoes the effect of assimilation rules that apply during phonological planning in production. Whether this is obtained through some kind of rule-based "reverse" phonology (Gaskell & Marslen-Wilson, 1996, 1998, 2001), or through a statistically based recurrent connectionist model (Gaskell et al., 1995; Gaskell, 2003), the principle is the same (even though processing issues are quite different). Such language-specific phonological inference mechanisms can account for the experimental results found with complete assimilation tokens presented above (Coenen et al., 2001; Gaskell & Marslen-Wilson, 1996, 1998). Crucially, they also predict that the pattern of compensation should depend on the listener's language.

Several studies have been investigating the perception of assimilated forms in a variety of languages, such as English (Gaskell & Marslen-Wilson, 1996, 1998), Dutch (Koster, 1987 ; Quené, van Rossum & van Wijck (1998), Japanese (Otake, Yoneyama, Cutler, & van der Lugt, 1996), German (Coenen et al., 2001; Weber, 2001), Hungarian (Mitterer et al., 2003) and French (Hallé, Chéreau, & Segui, 2000; Rigault, 1967; Snoeren, Hallé & Segui, 2006). Up to now, a few of them (Mitterer et al., Otake et al., Weber) present evidence in favor of such language-specific effects. However, they include a cross linguistic design in which listeners are presented with non-native phonology or ill-formed sequences. These results are therefore contingent on the problem of non-native speech perception and/or of phonotactic violations. In Mitterer et al. (2003), Hungarian and Dutch listeners had to identify the Hungarian word /bal/ 'left', which can be realized with a final [r] (rather as [ba$^{l}_{r}$] with a complex articulation) when concatenated to the suffix [ro:l] 'from the' (i.e. [barro:l]), but only as [bal] before the suffix [na:l] 'at the'. Therefore, the realization [ba**r**na:l] is an inappropriate assimilation. The identification task involving compensation and access to a lexical representation produced context effects and language-specific effects: Hungarian listeners had an identification bias towards the canonical [bal]-form when hearing the viable assimilation [barro:l]. This bias was absent in Dutch listeners, who were unable to *identify* (i.e. to decide whether they hear [bal] or [bar]) the syllables in the viable context – without ([balro:l]), or with assimilation ([barro:l]). However, clear conclusions are difficult due to the fact that these non-native listeners are hearing both nonwords and non-native phonemes. This result could thus be due to a more difficult discrimination, as shown by the authors. Indeed, they found an important difference between identification and discrimination tasks. For both groups, discrimination is more difficult in viable, than unviable contexts, and showed no effect of native language, indicating that it might be performed on the basis of lower-level, universal representations. When engaged in identification tasks, Dutch listeners don't seem to make use of the phonetic information given through the complex articulation in the stimuli, which would enable them to compensate for the change as do Hungarians. The authors conclude that *identification* performance seems to be influenced by language-specific experience (p.2323).

Other cross-linguistic evidence comes from Otake et al. (1996), showing that Japanese, but not Dutch listeners, were able to use nasal place assimilation in Japanese words (e.g. in *tonbo* 'dragonfly', where /n/ is realized as [m] vs. *konto* 'tale', with a dental [n]) to predict the post-assimilation context. This was the case despite the fact that the process tested (place assimilation in nasals) is present both in Japanese and in Dutch phonology (being optional in Dutch and obligatory in Japanese). Interestingly, Koster (1987) found that Dutch listeners were able to detect "a word ending in /n/" in assimilated [mb] sequences, but slower and with more errors than when it had no assimilation (*groe*[m] *boek*, vs. *groen book*). In this experiment (Koster, 1987, p. 98 – 102), words were produced with "complete neutralization", and half of the targets were having a lexical counterpart (*lijn* – *lijm* are both words), half were not (*groen* but **groem*). For Dutch listeners, therefore, a change from [n] to [m] is neutralizing and potentially blurs a lexical distinction. In Japanese, moraic nasals are never contrasting with respect to place of articulation, there is no possible word **komto* in Japanese (only non moraic nasals are contrasting in place of articulation, *tamago* 'egg' vs. *tanuki* 'rakoon' or *tanako* 'tenant'). The difference in behavior between Dutch and Japanese listeners may be due to the fact that Dutch listeners are hearing both non words and a different phonetic system, while Japanese might show compensation because this kind of assimilation in Japanese is obligatory and therefore, the canonical underlying representation itself might reflect assimilation. Again, like for Mitterer et al. (2003), conclusions are subject to the interpretation that Dutch listeners may not be able to perceive the moraic nasals in the same way as Japanese listeners do.

In Weber's study (2001), phoneme monitoring for the German fricative /x/ was used to test whether non-native listening is influenced when the non-native input violates a native assimilation rule (fricative assimilation in German (*la*[x]*t* 'laugh' vs. *li*[ç]*t* 'light'), being violated in Dutch non-word stimuli, e.g. [lixt]). Results showed that German, but not Dutch, listeners responded with a pop-out effect to violation of the German fricative assimilation rule. This effect is visible with non-native input though: the stimuli were recorded by a Dutch native speaker, and "sounded Dutch" (Weber, 2001, p. 101). In experiment 3 and 4, the design avoided the problem of presenting non-native input, but stimuli still contained a violation in the domain of phonotactics, where assimilation is obligatory in German (fricative assimilation and regressive nasal assimilation within syllables). Her results are therefore not directly informative with respect to the processing of legal native sequences.

So far, evidence for language specific listening has been obtained mainly through presenting non-native input to participants. In these conditions, such differences could also be due to violations of phonotactic constraints, or to unfamiliar sound categories, or even to syllable structure, in short, they are contingent on the problem of non-native speech perception. Therefore, the question remains, at least in the case of compensation for assimilation, whether processing of *legal* sequences in a *native* phonology is also dependent on phonological knowledge, or whether any change potentially reflecting assimilation would give rise to language-independent compensation effects (as suggested by Gow's results, 2001). In this sense, clear evidence in favor of language specificity in processing native input is rather sparse. In sum, all these results indicate some language specific elements in the processing of assimilated sequences, but do not give enough information about the way a possible model of word recognition would deal with assimilated words in a native language.

The present study

In order to further refine our understanding of language-specificity in compensation for assimilation, we designed a series of experiments, using a cross linguistic design but avoiding the problem of non-native speech perception. We included within the same language a native process as well as a non-native one, using exclusively the native categories of the listeners. We chose two comparable processes: regressive voicing and place assimilation. The first one exists in French, but not in English, whereas the second one exists in English, but not in French. Nevertheless both processes potentially neutralize phonemic contrasts of both languages. We therefore constructed French sentences containing occurrences of voicing assimilation (the native process) as well as occurrences of place assimilation (a non-native process). The same was done for English sentences.

In our experiments, listeners are processing only native speech, legal sequences and native phonetic categories in both conditions (place and voicing). Therefore any difference in compensation pattern that might emerge between the two conditions is hypothesized to reflect the use of language-specific knowledge of the process involved, rather than to differences arising from non-native speech processing.

As did most previous experiments on compensation for assimilation, we also considered context effects: occurrences of assimilation in our stimuli are either appropriate (i.e. surfacing in a suitable context for assimilation) or inappropriate (i.e. the context is normally not a trigger for the modification). We then distinguish two dimensions of modification in our stimuli: the native vs. non-native type of process, and within each, the appropriate (viable) vs. the inappropriate (unviable) context for the change. We also included a baseline condition in which the target word surfaced without any change, to ensure that in this case, detection is robust. Table 2 summarizes these experimental conditions:

Table 2. Experimental conditions for each type of process (native vs. non-native). Examples given for English stimuli

Condition	Place (native)	Voicing (non-native)
viable	we[p] pants	bla[g] glove
unviable	we[p] socks	bla[g] rag
no-change	wet shoes	black rug

The task we use is word detection: this is similar to identification, except that the actual response of the subject is a “similarity interpretation” rather than a “choice between two forms”: targets words are presented auditorily and followed by a sentence containing the target. But in the sentences, the targeted word surfaces either with a change (viable or not) or without any change (baseline). Participants are requested to press a button when they think that the target presented is the same in the sentence. A *yes* response then indicates that the word in the sentence is treated as a token of the target. A *no* response indicates that the change altering the word blocks its interpretation as a token of the target. This design then permits to obtain a measure of the degree of tolerance for modifications altering word forms. This is what we understand as *compensation*, i.e. when a change is compensated for, undone, in order to recover the “original/canonical” form of the word. If we see a difference in compensation between the native and the non-native type of change, this would be evidence in favor of the use of a knowledge of phonological processes during word recognition. In Experiment 1, French listeners are hearing French sentences, in Experiment 2, American English listeners hear American English sentences.

Experiment 1

Method

Stimuli

Thirty-two target items were selected. They were all monosyllabic nouns, with a C(C)VC structure. The target items consisted of two sets of 16 items: the Voicing Set and the Place Set, that were matched in average frequency (Place: 4238; Voicing: 4837, $t(15) = -0.4$, $p > .1$) according to the Brulex Corpus (frequency per 100 millions, from Content, Mousty & Radeau 1990, see the complete list of items in Appendix I). In the Voicing Set, all items ended in a final obstruent, that was voiced for half of the items, and unvoiced for the other half. Sixteen matched nonwords ([nw]) were constructed by switching the voicing feature of the final obstruents (e.g. *robe* /rɔb/ 'dress' - *rope* /rɔp/ [nw], or *lac* /lak/ 'lake' - *lague* /lag/ [nw]). In the Place Set, final consonants were all coronal; half were nasals and half were stops. Sixteen matched nonwords were obtained by a change in the place feature (12 towards labial, 4 towards velar) of the final consonant (e.g. *moine* /mwan/ 'monk' - *moime* /mwam/ [nw] or *guide* /gid/ 'guide' - *guibe* /gib/ [nw]). Each of the 32 target items was associated with a triplet of context words. In French, context words were always adjectives since the standard noun phrase has the shape 'determiner noun adjective'. Each adjective in a triplet corresponded to one of the experimental conditions: *viable change*, *unviable change*, and *no-change*. For the viable change condition, the adjective's initial consonant was an obstruent agreeing in voicing or in place with the nonword matched to the target item, depending on the item set (e.g. in the Voicing Set: *rope sale* /rɔ**ps**al/ 'dirty dress$_{[nw]}$'[1]; in the Place Set: *moime bavard* /mwa**mb**avaʀ/ 'talkative monk$_{[nw]}$', respectively. The adjectives in the unviable change and no-change conditions both started with a neutral consonant which was not involved in the relevant assimilation process. For the voicing set, this neutral consonant was always a sonorant (nasals and liquids, as well as the standard French uvular fricative [ʁ]), that does not trigger voicing assimilation in French. In the Place Set, this neutral consonant was either a sonorant, a coronal or labiodental fricative, or the coronal stop [d]; none of these consonants is involved in place assimilation in English. In all 3 conditions of both the Voicing and the Place set, the association (pseudo)noun-adjective always yielded a legal consonant cluster in French and did not contain any violation of voicing or place assimilation[2].

Finally, 3 sentence frames were constructed for each of the 32 target items. A sentence frame consisted in a sentence beginning and sentence ending, where each of the three (pseudo)noun-adjective combinations could be inserted and resulted in a plausible sentence (e.g. *Elle a mis sa* ___ ___ *aujourd'hui.* 'She put on her ___ ___ today.'). Globally, the sentence frames were matched in number of words and position of the insertion slots across the Voicing Set and the Place Set. No occurrence of violation of voicing or place agreement occurred in the frames neither. Combining the three conditions with the three sentence frames gave rise to 9 actual sentences associated to each item. This resulted in a total of 288 sentences.

For purposes of counterbalancing, we defined three experimental lists. In each list, all three conditions were present for each item, but in different sentence frames. The sentence frames were rotated across

[1] Here, [nw] means that the word underwent an assimilatory change, and became a nonword.

[2] This constraint made it necessary to include geminate clusters in the place set, otherwise the place agreement would have also produced violation of the voicing agreement constraint in French. In order to balance both sets, we also included the same number of geminates in the voicing set. The speaker produced all geminates as a single long consonant, without release in between. The same constraint has been obeyed for English stimuli sets.

the three lists, so that across the experimental lists all three conditions appeared in all three sentence frames. Thirty additional filler sentences were constructed that were similar to the experimental sentences (same kind of alterations on the target involving one feature, same proportion of identical and changed words), and served as training (N=18), or distractors (N=12). Modifications involved voicing, manner and place contrasts at the end or beginning of target words, in order to drive participant's attention to the precise form of words (e.g. target "cube", filler sentence containing "gube" [gyb]). Crucially, these filler sentences did not contain any case of assimilation in either viable or unviable context, so that the feedback provided here was unambiguous and could not influence later participant's responses on test sentences.

The 288 test, 12 distractor and 18 training sentences were recorded by the first author, a female native speaker of French[3]. They were digitized at 16kHz and 16bits on an OROSAU22 sound board, and edited using the sound preparation software CoolEdit and Praat. The onset of the carrier word and the onset of the following adjective were marked through digital labels. The 32 target words for the experimental sentences and 30 targets for filler sentences were recorded by a male native speaker of French and digitized.

Procedure

This experiment was run using the *Expe6* stimuli presentation program (Pallier, Dupoux, & Jeannin, 1997; http:\\www.lscp.net\expe). The experimental trials consists in the presentation of the target item (male voice), followed after 500ms of silence by a sentence (female voice). Participants are requested to press a button when they think that the target presented is the same in the sentence, and refrain from pressing otherwise. This instruction – together with the specific training – was given in order to draw their attention on the detail of pronunciation of words, i.e. on the *form* of words and not to the mere presence or absence of a target word in the sentence. For the same reason only a few distractor sentences were included. This instruction was important in order to make participants understand that they have to be precise in their judgments and not only press *yes* if they recognized semantically the target word in the sentence. Otherwise, such minimal differences would have been at risk to be ignored in a word detection task. Several studies (McClelland & Elman, 1986; Norris 1994) show that a word is still recognizable even if changes altered its canonical form. The degree of "recognizability" is inversely proportional to the word's frequency and neighborhood density. We therefore chose frequent monosyllabic words, in order to augment the importance of any minimal change affecting the word form. Participants are told to respond as quick as possible, without waiting until the end of the sentence. They were allowed in total 3000ms after the word onset (in the sentence) to make their

[3] For this and the following experiments, all speakers were trained until they are familiar with the nonwords, and able to pronounce all sentences in a natural way. We avoided cross splicing due to the difficulties to match whole sentences with respect to prosody and speech rate.

response. After that delay, the next trial is initiated. During the training phase (18 sentences), feedback was provided whenever the participants gave an incorrect response, that is, failed to detect the target word or incorrectly pressed a button for a non-target (the training sentences did not contain any occurrence of viable or unviable context). During the test phase, responses were collected without feedback. The test phase was split into three blocks of 36 trials that were constructed such that a given test item appeared only once within each block. A pause was inserted after each block to allow participants to rest and concentrate. Order of trials within each block was separately randomized for each participant. The experiment lasted 20 minutes. Instructions appeared on the computer screen, and were completed orally by the experimenter when needed.

Participants

Eighteen French native speakers (all grew up monolingually, having only limited and late experience with English) were tested on this experiment, individually and in a quiet room. There were 11 women and 7 men, all living in the Parisian area. They ranged in age from 19 to 28 years. None of them took part in a similar experiment previously, and none of them reported any history of hearing impairment. They were randomly assigned to one of the three experimental lists. They were paid for participation.

We expected participants to detect the target words in the no-change condition, and to reject them in the unviable change condition (given previous results from the literature). The performance on these two conditions serves as a baseline to evaluate the responses in the viable change condition. If participants fully compensate for the phonological rule, they should detect the target word to the same extent as in the no-change condition, despite the fact that the target underwent the same featural change as in the unviable change condition. If there is no compensation for the phonological rule, participants should respond like in the unviable change condition, that is, reject the changed word as a non-target. Hence the logic of our experiment is similar to that of Gaskell and Marslen-Wilson (1996, 1998), except that we use word detection instead of cross-modal priming or phoneme detection.

Control task: forced-choice judgment on spliced-out target words.

To ensure that the critical items' final consonants were unambiguously perceived as changed or unchanged, we first carried out a pretest in which we excised all target words out of the carrier sentences and presented them in isolation in a forced-choice categorization task. Words were presented auditorily and followed by a 3 s. silence, during which participants had to tick the consonant they heard on a response sheet. They always were given a choice between the original consonant and the assimilated one. For the word *robe* 'dress' for example, the choice was between [b] (unchanged) and [p] (underwent voice assimilation). A free cell allowed them to report any better matching sound, if needed. The entire procedure lasted about 18 minutes. Eighteen French native speakers who had not participated in the other study were recruited to take part in this control experiment.

Results

We report first the results from the pretest, summarized in Table 3.

Standard error (SE) is given in parentheses. Results include the whole data set (all items and participants).

Table 3. different consonant judgment rate (%) across contrast type and condition for French stimuli (n=18)

	Consonant different from unchanged target (%) :	
	Place (SE)	Voicing (SE)
viable change	92 (0.9)	95 (0.7)
unviable change	90 (1)	97 (0.5)
no-change	9 (2)	2 (0.2)

This table shows clearly that both change conditions yield in majority "different consonant" responses, there is no significant difference between both change conditions (an ANOVA with subjects as random variable, restricted to both change conditions for place and voicing together, yielded no effect of condition ($F(1,17) = 0.2$, $p>.6$). Items in the no-change condition are judged largely as having a "similar consonant" (to 91% and 98%). Globally, contrast type has no effect either ($F(1,17)=4.2$, $p>.05$).

For the word detection task, we checked whether some items triggered too many errors in the baseline conditions, namely the no-change and unviable change conditions. All items that yielded detection values higher than 50% in the unviable change condition (i.e. more than 50% false alarms) or less than 50% in the no-change condition (i.e. more than 50% misses) were excluded. One Voicing item was dropped (*badge*).

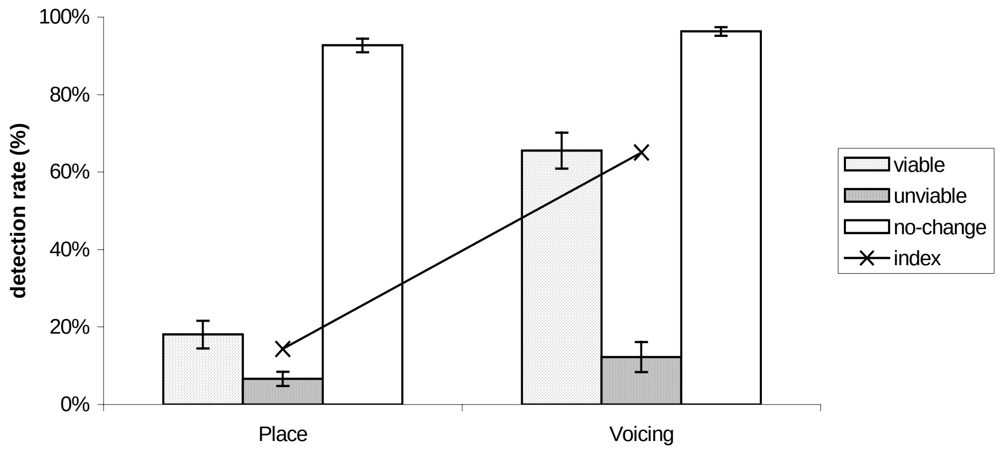

Figure 1: French listeners, French sentences: Detection rate in each condition, for both place and voicing assimilation types, N = 18

The percent detection rate was subjected to two analyses of variance (henceforth ANOVA), one with participants, one with items as random variable. The by-subjects ANOVA had one between-subjects factor, Group (counterbalancing factor, 1, 2 or 3) and two within-subject factors, Condition (*viable change*, *unviable change* or *no-change*) and Contrast (*Voicing* or *Place*). The by-items ANOVA had one between-item factor, Contrast and one within-item factor, Condition. We observed a main effect of Condition ($F_1[2,30]=635.8$, $p<.0001$; $F_2[2,58]=448$, $p<.0001$), a main effect of Contrast ($F_1[1,15]=63.8$, $p<.0001$; $F_2[1,29]=54$, $p<.0001$), as well as an interaction between these two factors ($F_1[2,30]=55.2$, $p<.0001$; $F_2[2,58]=37.1$, $p<.0001$), suggesting that the two item sets behaved differently across the three conditions. The Group factor showed no main effect and did not interact with the other two factors. Similarly, the same analyses declaring the factor Blocks (1, 2 or 3) instead of Group revealed that there were no effects of Blocks in subjects or items, suggesting that repeated presentation of the same word targets across different blocks did not cause any benefit or cost in processing. Mean detection rates are displayed in Figure 1 as a function of Contrast and Condition. Examination of mean detection rates revealed that the difference between the Voicing and the Place set was mainly in the viable change condition (65% for the Voicing Contrast vs. 18% for the Place Contrast, effect size 47%, $F_1[1,17]= 72.4$, $p<.0001$; $F_2[1,29]=58.7$, $p<.0001$). In contrast, the other two conditions behaved similarly for both contrasts (14% vs. 06% in the *unviable* condition, effect size 8%, $F_1[1,17]= 2.1$, $p>.1$; $F_2[1,29]=2.9$, $p=.094$; 96% vs. 92% in the *no-change* condition, effect size 4%, $F_1[1,17]= 4.4$, $p=.05$; $F_2[1,29]=3.2$, $p=.082$).

To further refine our analysis, we computed for each subject and item an index *x* of compensation (formula 1) on the basis of the number of yes-responses as a function of condition and contrast type (place vs. voicing) :

$$\text{(1) Compensation index} = \frac{(\text{detection}_{\text{viable change}} - \text{detection}_{\text{unviable change}})}{(\text{detection}_{\text{no-change}} - \text{detection}_{\text{unviable change}})}$$

This index calculates the relative value of detection in the viable condition as a function of both other conditions. This allows to obtain the *ratio* of “viable” to “no-change”, controlling for response biases or errors from the “unviable” condition. The index *x* thus corresponds to the *degree* of compensation for either place or voicing type of change. If participants fully compensate for assimilation, they will detect the target word in the viable change condition as often as in the no-change condition: the index will be 1 (since the numerator and the denominator will be equal). If participants do not compensate at all for assimilation, they will respond to the target in the viable change condition as rarely as in the unviable change condition: the index will be 0 (since the numerator will be 0).Values of the index intermediate between zero and one will indicate partial compensation for assimilation.

We computed the compensation index for each participant and each contrast (mean index for participants is 0.65 (65%) for voicing and 0.14 (14%) for place), and used it as the dependent variable in an ANOVA with Contrast as a within-subject (respectively between-items) factor. We found a significant effect of Contrast, with a higher index of compensation for Voicing than for Place, confirming the fact that participants compensate significantly more for voice assimilation than place assimilation (65% vs. 14%, effect size 51%, $F_1[1,17]=77.4$, $p<.0001$; $F_2[1,29]=51.2$, $p<.0001$).

Reaction times were collected for a “yes response”. Mean times by subjects are comprised between 519 ms and 2107 ms (mean RT for n=18: 1582 ms). The experiment was fairly speeded: the time to make a response was limited, and participants should not wait until the end of the sentence. Overall, it should be noted that this experiment is demanding, speech rate is fast and contrasts are minimal. The slow RTs we got surely do not completely rule out the possibility of strategic responding. But we did our best to limit the risk of such a response pattern in our participants. A concern about offline strategic responding can however be reasonably rejected, as post-hoc analyses revealed no difference about the pattern of results according to slow vs. fast reaction times (ANOVA by subjects including the factor *RT* (fast vs. slow) and the factors *condition* and *type* revealed no interaction of the RT factor with both other factors).

Discussion

Experiment 1 revealed two main results. First, French participants compensate for voicing assimilation in a context-sensitive fashion: viable contexts give rise to higher detection rates than unviable contexts. These results show a context effect comparable to the one observed by Gaskell and Marslen-Wilson (1998) with English listeners for a native assimilation process in English: place assimilation. We were also able to show that this compensation was not complete, however, since the compensation index only reached 65% (and was significantly different from 100%). This suggests that complete assimilation may not be the most natural case in French and that the word recognition processor is only able to compensate partially for such extreme cases. An alternative explanation could be that participants perform this recognition task integrating information from different processing levels simultaneously (*multiple readout hypothesis*, similar to Grainger & Jacobs, 1996, or to the Race Model, Cutler & Norris, 1979): the phonological level, representing a phonological form (recovered or not by a compensation mechanism), the lexical level, and a language independent phonetic level. A similar hypothesis (the dual task) has been evoked by Gaskell and Marslen-Wilson (1998), who observed that detection of phonemes in real words was higher than in nonwords. In our experiment, intermediate compensation (65%) may be the product of combining information from all levels: Faced with a (minimally deviant) word form, the lexical level leads to a “yes” response. The phonological level reinforces a “yes” response when the change is viable or has been compensated, whereas the phonetic form detector yields a “no” response.

The second main result from Experiment 1 is that French participants compensate much less for place assimilation, a rule that does not exist in French (the compensation index is only 14%), than for voicing assimilation. Since Gaskell and Marslen-Wilson (1998) previously obtained sizable compensation for place assimilation with British English participants and sentences (60% /t/-detection in assimilated *freigh[p b]earer*), this result corroborates that phonological compensation is language-specific. We will come back to this point in Experiment 2.

French participants nevertheless did compensate somewhat for place assimilation: even though the Place change does not correspond to an existing rule in French, participants treated 18% of the words appearing in the viable change condition as tokens of the target as opposed to only 6% of the words in the unviable change condition. The presence of a (small) context effect for this contrast (index value is 14%) suggests the existence of a language independent compensation mechanism in addition to the language-specific one; it nevertheless seems to be the case that the universal mechanism has a weak influence compared to the language-specific one, at least in a task involving complete changes. We are currently investigating whether this result reflects a general preference for homorganic consonant clusters, related for example to the high frequency of place assimilation phenomena across the world's languages.

So far, the difference observed in compensation between native and non-native assimilation suggests that compensation for assimilation reflects a phonological knowledge of these processes: This conclusion stems from the fact that French speakers showed greater compensation for voicing assimilation (a native rule), than for place assimilation (a non-native rule). However, this single experiment can not exclude the possibility that independent phonetic differences between voicing and place induced the results (see discussion section in Experiment 2). Indeed, it could be that voicing cues are intrinsically weaker than place cues in the context tested (VC#CV clusters), thus allowing for an easier acceptance of changed forms as being "the same", i.e. inducing more "compensation" before other obstruents which mask the preceding consonant. It could then happen that native listeners of other languages too would compensate more for voicing than place assimilation, whatever the rules actually present in their native language. At first sight, however, it seems not to be the case that voicing cues are intrinsically weaker than place cues. Indeed, voicing is a quite robust cue for several reasons: first, voicing is periodic in nature, distributed over lower regions of the spectrum than place, making it more robust to noise (Wright, 2004). Second, because different acoustic parameters are involved (vowel duration, duration of voiced portion in closure, closure duration, VOT-lag, F0,…) which all contribute to the voicing distinction (see Kohler, 1984; Kingston & Diehl, 1994, among others), listeners probably have more converging cues to this contrast. Indeed, place cues for stops are said to be weaker especially in this word-final cluster environment (VC#CV), where release burst is not reliable. Place cue markers are therefore restricted to VC-formant transitions, and are more variable in this VC position than in the CV position (Wright, 2004; Jun, 2004, p. 61). Because these are periodic as well, though, they resist quite well to masking, especially in optimal listening environments. An independent reason for considering voicing as being equal to place with regard to clarity is that the results of the control experiment did not show increased error rate for voicing items as compared to place, what would have been the case if voicing cues were less perceptible than place cues.

The possibility that place and voicing cues differ in strength in this environment seems implausible, and therefore we tend to interpret the results of the French listeners as support for a language specific compensation mechanism. However, in order to establish more strongly that compensation reflects language-specific knowledge of processes, and not only the language-independent use of phonetic properties, we need to test English participants with the same experimental design as we used for French participants. We expect the English participants to behave differently from the French participants: they should compensate more for Place than for Voicing assimilation. In contrast, if compensation for assimilation is largely language independent and based on differences between voicing and place, then English participants would behave much like French participants, and compensate more for Voicing than for Place assimilation.

English has no voicing assimilation rule, but a rule of place assimilation affecting coronal stops. Experiment 2 involves American English participants.

Experiment 2

Method

Stimuli

Following the same method used for French stimuli, thirty-two target items were selected. They were all monosyllabic adjectives, with a C(C)V(C)C structure. Target items were split into two sets of 16 items: the Voicing Set and the Place Set. They did not differ in average frequency (per million, according to both the *Phondic* Database, and the Kucera & Francis Word Frequency as given in the MRC Psycholinguistic Database (Wilson 1988) : voicing: 151 (K&F: 144), place: 156 (K&F: 152), $t(15)=.06$, $p>.1$; see the complete list of items in Appendix II). In the Voicing Set, all items ended in a final obstruent, that was voiced for half of the items, and unvoiced for the other half. Sixteen matched nonwords ([nw]) were constructed by switching the voicing feature of the final obstruents (e.g. /nais/ (nice) - /naiz/ [nw], or /bik/ (big)-/bik/ [nw]). In the Place Set, all final consonants were coronals, and half were stops, half were nasals. Sixteen matched nonwords were obtained by a change in the place feature (towards labial or velar) of the final consonant (e.g. /swiːt/ (sweet) - /swiːk/ [nw] or /pleɪn/ (plain) - /pleɪm/ [nw]).

Each of the 32 target items was associated with a triplet of context words; In English context words were always nouns because the standard noun phrase in English is 'determiner adjective noun'. Each noun in a triplet corresponded to one of the experimental conditions as defined in Experiment 1: viable change condition, unviable change condition, and no-change condition. For the viable change condition, adjectives started with an obstruent agreeing with the nonword matched to the target item; the nature of agreement was the same as described for Experiment 1 (place, e.g. [fæp pʌpi] 'fat$_{[nw]}$ puppy' or voicing, e.g. [blæg glʌv] 'black$_{[nw]}$ glove'). Nouns in unviable change and no-change conditions for the Voicing Set started with a nasal or a liquid, consonants which are not involved in a voicing assimilation process. In the Place Set, nouns in both unviable change and no-change conditions started preferably with coronal sonorants, sometimes with coronal fricatives or the coronal stop [d] (the proportion of sonorants to obstruents is 5 to 3 in the place-stop list, and 2 to 6 in the place-nasal list). None of these consonants is involved in place assimilation processes in English. For the unviable change condition, the noun would be associated to the nonword matched with the target word (e.g. [blæg ɹæg] 'black$_{[nw]}$ rag'). In the no-change condition, it would be associated to the target word itself (e.g. [blæk ɹʌg] 'black rug'). In all 3 conditions, the association (pseudo)adjectives-noun always yielded a legal cluster in English. There were no coronal-labial or coronal-velar clusters, in order to avoid spurious effects due to violation of the place assimilation rule.

Finally, 3 sentence frames were constructed for each of the 32 target items following the same method as used for French sentences. This resulted in a total of 288 sentences. Three experimental lists were defined similarly to those used in Experiment 1.

The 288 test, 12 distractor and 18 training sentences were recorded by the fourth author, a female native speaker of American English (her speech corresponding to General American Standard), living in New Haven, CT. Target words were recorded by a male native speaker of American English from New York. They were digitized at 16kHz and 16bits on an OROSAU22 sound board, and edited using the sound preparation software CoolEdit and Praat. Onsets of the carrier words and onsets of the following adjectives were marked through digital labels.

Procedure

The same procedure was used for the presentation of the stimuli. However, we used the *E-prime* stimuli presentation program (www.pstnet.com/e-prime/default.htm) instead of *Expe6*, due to hardware reasons. We also slightly modified the instructions: Participants had to press a "yes" button when they thought that the target was present in the sentence, and a "no" button otherwise.

Participants

Twenty-six Americans aged from 18 to 53, from the North-East of the U.S. (mainly New England), were tested on this experiment in Paris (France), in Providence (RI), New Haven (CT) and Amherst (MA). They all grew up monolingually, and came roughly from the triangle between Washington DC in the south, Chicago in the West and Boston in the North-East. None of them took part in a similar experiment previously and none of them reported any auditory deficits. They were paid for participating. All of them had late experience with French, 19 of them were living in France by the time of testing. They were tested on French sentences in the same testing session, half of them before American English, half of them afterwards. Nine participants were highly fluent in French, the 17 remaining were beginning learners. Their results on French sentences are presented in Darcy, Peperkamp and Dupoux (in press).

Control task: forced-choice judgment on spliced-out target words.

As in Exp. 1, all target words were excised out of the carrier sentences and presented in isolation in a forced-choice categorization task. Sixteen American native speakers who did not participate in any of the previous studies were recruited to take part in this control experiment.

Results

Table 4 presents the results of the forced-choice categorization task. Results include the whole data set (all items and participants).

Table 4. different consonant judgment rate (%) across contrast type and condition for American English stimuli (n=14)

	Consonant different from unchanged target (%) :	
	Place (SD)	Voicing (SD)
viable change	74 (3)	78 (1)
unviable change	78 (2)	77 (1)
no-change	23 (4)	17 (3)

As can be seen from Table 4, both change conditions yield an equal amount of "different consonant" responses, there is no significant difference between both change conditions (an ANOVA with subjects as random variable, restricted to both change conditions for place and voicing together, yielded no effect of condition ($F(1,13) = 2.3$, $p>.1$). Items in the no-change condition are judged largely as having a "similar consonant" (to 80% on average). Globally, contrast type has no effect either ($F(1,13)=0.1$, $p>.6$).

One striking difference compared to the French results (see Table 3) is the higher error rate visible in the American English categorization results. However, this difference is not central to our argument. The most critical result to be seen in both control experiments is the absence of any difference in the "clarity of changes" between place and voicing targets, given the suggestion made above that voicing may have less clear cues, therefore favoring compensation over place targets. For both experiments, the answer is "no": in isolation, cues seem to be equal for voicing and place targets, and can not explain any observed differences in behavior. We return to the question of higher error rate in the discussion section for Experiment 2.

Using the same criterion for item rejection as in Experiment 1, 4 items were rejected, 1 in the Voicing set, 3 in the Place set.

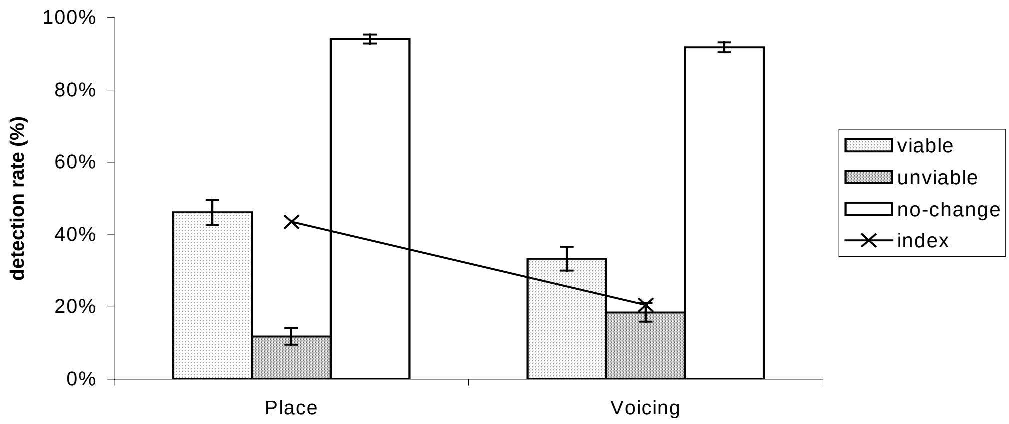

Figure 3: American listeners, American English sentences: Detection rate in each condition, for both place and voicing assimilation types, N = 26

Mean detection rate was subjected to two ANOVAs, one with participants, one with items as random variable. The participants ANOVA declares the between-subject factor Group (1, 2 or 3), and two within-subjects factors: Contrast (*place vs. voicing*) and Condition (*viable change vs unviable change*). As above, the by item ANOVA declared one between item factor Contrast and one within-item factor, Condition. In the participant analysis, no effects related to the factor Group became visible. We observed a main effect of Condition ($F_1[2,46]=468.9$, $p<.0001$; $F_2[2,52]=181.9$, $p<.0001$). The Contrast effect was almost significant by participants, but not by items ($F_1[1,23]=3.5$, $p=.07$; $F_2[1,26]=0.3$, $p>.1$). We found an interaction between these two factors that was significant only by participants, marginal by items ($F_1[2,46]= 40.2$, $p<.0001$; $F_2[2,52]=2.7$, $p=.07$), evidencing that they behave differently according to the contrast type (place vs. voicing) across conditions. Items display more variability, to which we will return below. Mean detection rates as a function of contrast and condition are displayed in Figure 3. The viable change condition yielded 33% detection responses for the Voicing Contrast, and 46% for the Place Contrast, a significant difference by participants (effect size 13%, $F_1[1,25]=32$, $p<.0001$; $F_2[1,26]=1.7$, $p>.1$). The no-change condition was very similar in both contrasts (94% detection for Place vs. 91% for Voicing, effect size 3%, $F_1[1,25]=1.8$, $p>.1$; $F_2[1,26]=0.6$, $p>.1$). Detection rate in the unviable change condition was different between the Place and the Voicing contrast, significantly only by participants (11% vs. 18% for each contrast respectively, effect size 7%, $F_1[1,25]=11.4$, $p<.01$; $F_2[1,26]=1.3$, $p>.1$).

We computed the compensation index according to formula (1) for each participant and each item (mean index is 20% for voicing and 43% for place), and used it as a dependent variable in an ANOVA first by participants, then by items. We declared Contrast as a within-subject (respectively between-item) factor (Place versus Voicing). We found a significant effect of Contrast by participants (not by items), confirming that all subjects behave similarly and compensated significantly more for place assimilation than voicing assimilation ($F_1[1,25]=57$, $p<.0001$; $F_2[1,26]=2.7$, $p>.1$). A t-test revealed that compensation for assimilation was not complete in the Place condition, since the compensation index was significantly different from 100% ($t_1(25)=14.6$, $p<.0001$; $t_2(12)=7.6$, $p<.0001$). For the Voicing contrast, the index differed significantly from zero ($t_1(25)=5.7$, $p<.0001$; $t_2(14)=2.6$, $p<.05$).

Mean reaction times by subjects are comprised between 1285 ms and 2485 ms (mean RT for n=26: 1920 ms). Analyses of reaction times and detection values did not reveal any interaction of *RT* with the factors *condition* and *type*.

In this experiment, variability in items inhibited various significant effects in our analyses. Looking in greater detail at the pattern of this variability, we see that it mainly concerns voicing items. Place items behave homogenously. Voicing items display an asymmetry between voicing and devoicing items (e.g. *tough* vs. *big*). Compensation was higher for devoicing items: this means that detection (compensation) is higher for 'big fountain' *bi[kf]ountain* (34%) than for 'tough demand' *tou[vd]emand* (08%). The difference between indices for voicing vs. devoicing is significant by participants and items ($F_1[1,25]=23.5$, $p<.0001$; $F_2[1,13]=5.6$, $p=.03$). This could reflect compensation for a process of partial phonetic final devoicing applying in American English (Hyman 1975, Keating 1984, p. 293). Therefore, for Americans, only the *voicing* items are really non-native. When restricting the analysis to those items, the difference between indices for place and voicing (without *devoicing* items) is very significant by subjects and by items ($F_1[1,25]=34.5$, $p<.0001$; $F_2[1,19]=8.8$, $p<.008$).

Pooled analysis with both experiments on detection rates was performed in order to examine whether listeners' behavior is different across languages, and whether the factor test-language interacts with differences due to contrast type or to condition. Mean detection rate was subjected to a ANOVA with participants as random variable. We declare the factor "test-language" (French or English), as well as both crucial factors "condition" and "contrast". The factor "test-language" yields no significant main effect, because the directions of effects cancel each other out ($p>.7$). "Test-language" interacts strongly with "contrast" ($F_1[1,42]=54.4$, $p<.0001$) and in a triple interaction also with "condition"

($F_1[2,84]=91.4$, $p<.0001$). This means that both experiments show an opposite pattern of detection, where the test-language strongly influences detection according to contrast type as well as condition.

Discussion

The main result from Experiment 2 is that American participants listening to American English sentences showed a pattern of results symmetrical to the one observed for French participants listening to French. This result clearly supports the hypothesis that compensation procedures are partly governed by language-specific phonological knowledge. More precisely, we observed that American listeners compensated significantly for changes that correspond to the application of the place assimilation rule in American English. They also compensated for voicing, a process which is not native. However, further analysis of compensation differences between voicing and devoicing revealed that it might be necessary to consider “devoicing” as a native process rather than a non-native one, as opposed to “voicing”, which can definitely be considered as “non-native”, and for which compensation is considerably reduced. In sum, the difference observed in compensation patterns between place and voicing provides further support for the assumption that compensation is driven by language-specific knowledge of phonological processes.

There is one important difference, though, between the French and the American experiments: the amount of compensation for the native rule was larger in French than in American English (65% vs. 46%). This could be due to the fact that place assimilation is less systematic in English than voice assimilation is in French (see Otake et al., 1996, for a similar observation). In other words, the word recognition system for English listeners would be less used to cope with complete place assimilation, than it is used to with complete voicing assimilation in French. When a word is heard in a sentence context, compensation mechanisms are at work, and if they are presented with “optimal” stimuli for which they have been tuned for in the course of language acquisition, they are predicted to be most successful. In our case, the reality of English place assimilation makes our stimuli (because they present rather categorical changes) not optimal for the system to compensate for. This might be slightly different for French stimuli, if categorical changes we present parallel more closely the reality of French voicing assimilation the system is used to. One could argue that the difference in compensation rate between English and French could originate in the degree of variability in phonetic cues in our stimuli, being more variable in English than in French. Even if this might indeed be present in the stimuli, as indicated by the difference in error rates in the categorization experiment (see below), it does not explain the different compensation patterns in Experiment 1 and 2, for two reasons. First, in case compensation would be the mere reflect of tolerance to cue-uncertainty, one would expect more tolerance in the English case, where cues seem to be more variable, more ambiguous than in French. The difference, however, goes in the opposite direction. Second, one would not expect to find any difference due to condition between viable and unviable condition, i.e. the correct rejection in unviable context (context effects for the native process). For both experiments, the percentage of false alarms in this condition is similar and rather low: for French listeners, voicing yields 06% false alarms, for English listeners, place yields 11%, false alarms in the unviable context. The difference to the respective detection rates in viable conditions is striking (French 65%, English 46%).

The difference observed in the categorization results between English and French – where English listeners make more errors (around 20%) – could reflect a general tendency of phonetic cues to being more variable or less robust in English than in French, especially in this context (see discussion of Experiment 1). Numerous studies have shown systematic differences in the phonetic implementation of particular contrasts between French and English or other languages, with particular attention to the voicing distinction markers (Mack, 1982, Kohler 1981, among others). To our knowledge, no study so far examined such systematic differences in cue variability or robustness between English and French, in word-final position before obstruents. Some indirect evidence is found in cross-linguistic studies of intelligibility in time-compressed speech. For a similar compression rate of 50% in English and French sentences, English listeners are able to recall only 44% of the syllables, whereas French listeners listening to compressed French show recall-scores averaging 85% (Mehler et al. 1993, Sebastian-Gallés et al, 2000). In sum, there is a difference in the overall clarity of cues due to particularities of American English and the respective implementation of cues in the particular contexts used. But this cue-robustness difference does not explain the pattern of compensation found in Experiments 1 and 2.

General Discussion

The main goal of this study was to investigate the existence of a language-specific phonological knowledge involved in compensation for phonological assimilation. We conducted first three experiments, testing two different phonological processes on different languages. Experiment 1 investigated compensation in French native speakers on French stimuli: participants showed more compensation for the voicing contrast than for the place contrast, but only in viable contexts for French voicing assimilation. In Experiment 2, speakers of American English were tested on American English sentences using the same task: participants compensated more for the place contrast than for the voicing contrast, and only in viable contexts for English place assimilation, thereby presenting symmetrical results from Experiment 1. All these results are supported by additional control experiments, carried out to eliminate the possibility that results could be due to unintentional bias in the stimuli. Excised targets were presented in a forced choice task to new listeners of each language. Words in both change conditions for place and voicing equally were perceived as being different from the form of the target in isolation, meaning that changes were perceived clearly.

Therefore, higher detection rates visible in viable change conditions for the respective native processes is attributable to phonological compensation for assimilation, involving a language-specific knowledge of the processes at work in the language, rather than the language independent use of phonetic cues. Additional support for this view are the results presented in Darcy, Peperkamp and Dupoux (in press): In these experiments, the same listeners – L2 learners of the other language – were presented to both languages, French and American English. French listeners who were beginning learners of English showed the same behavior on both languages, compensating more for voicing assimilation than for place assimilation (69% vs. 40% in French, 64% vs. 37% in English, difference between voicing and place significant). Similarly, American English listeners, who were beginning learners of French (the same participants as in this Experiment 2), showed upon hearing French sentences the same pattern of compensation as they show here, hearing American English sentences (voicing vs. place: 32% vs. 49% in French, and 33% vs. 46% in American English). The fact that they do show a different pattern of compensation on the same stimuli as did the respective native speakers of that language is to be interpreted in the way that these learners still did not acquire the compensation mechanism for that specific process in L2. It excludes the possibility that the observed difference is the result of unintended bias in the stimuli, as here the manipulated variable is only the listener’s L1s.

These results converge in showing that compensation is not driven by the unintended acoustic differences between both languages, but rather by the phonological knowledge of the way assimilation works in one language.

Because lexical compensation mechanisms are not sensitive to phonological context, such mechanisms alone cannot explain our results. Similarly, phonetic compensation mechanisms do not rely on familiarity with specific phonological processes, and therefore cannot explain our results either. Nevertheless, we do not think that such mechanisms must necessarily be ruled out. In fact, our data are compatible with the existence of such mechanisms alongside a phonological language-specific, context-sensitive mechanism. The three types of mechanisms would operate at distinct levels of representation, and would all influence subjects’ responses in a given task.

To elaborate on our proposal, we postulate that beyond basic auditory processing, speech is initially represented in a universal phonetic format; at this level, language independent mechanisms such as feature parsing may operate (Gow, 2001, 2002a; Gow & Im, 2004; Gow & Zoll, 2002). At the next stage of processing, speech is encoded in a language-specific phonological format; at that level, language-specific mechanisms such as phonological inference to compensate for phonological alternations may operate (our data, Gaskell & Marslen-Wilson, 1996, 1998). Finally, such phonological representations are matched against lexical representations for word recognition, in the manner described by multiple activation models (Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978; McClelland & Elman, 1986; Norris, 1994). Behavioral responses can be influenced by any of these processing levels (as predicted by a multiple readout model). Which level has the greatest influence on behavioral responses depends on many factors, including the task (word identification vs. discrimination), and the nature of the stimuli: whole sentences vs. isolated words or syllables; words vs. nonwords; with large acoustic variations (e.g. across different speakers) or not.

Postulating multiple and cascading compensation mechanisms makes it possible to reinterpret apparently conflicting results from the literature. In the present experiments, we have maximized our chances of observing effects reflecting phonological processing by using words embedded in sentences, and identification across different speakers. Other studies that have used discrimination of nonwords produced by the same speaker have obviously maximized the influence of the phonetic processing level, thereby explaining their finding of universal patterns of compensation.

Gow (2002b) and Gow and Im (2004) reported language independent low-level effects of compensation for voicing assimilation in Hungarian, whether the subjects were native speakers or not (e.g., Korean listeners). These results seem in contradiction with ours. However, it should be noted that these studies used different stimuli from ours: Rather than presenting complete assimilations, they presented ambiguous (multiply articulated) segments, thereby favoring feature parsing. Furthermore, we would like to argue that detecting a word within a sentence across voice changes, the method we used, should force listeners to recode the stimuli at the phonological level and give greater weight to that level in the decision process, as fine acoustic/phonetic details are irrelevant and even interfere with this task. On the other hand, detecting phonemes within bi-syllables without much acoustic variation (their task) may well be more easily performed by paying attention to the phonetic level of representation. According to this interpretation, both our results and those of Gow and Gow and Im can be explained by the same multiple readout model; simply, their experiments induce responses predominantly based on phonetic representations and therefore reflect universal phonetic processes, whereas our experiments (and those of Gaskell & Marslen-Wilson, 1996, 1998) induce responses based primarily on phonological representations, therefore reflecting language-specific abstract phonological processes.

Restated within this framework, our results show that the phonological level is responsible for most of the effects observed in our experiments, as it is the only level where both context-sensitive and language-specific effects may arise. But even before this phonological inference mechanism applies, some degree of universal feature parsing may occur, prompted by e.g. homorganic clusters. This effect could explain the small, but non-null compensation for voicing assimilation by English listeners, and for place assimilation by French listeners. Finally, lexical compensation mechanisms may also have played a role in our experiments. Such a mechanism would generate a global tendency to detect the target based on phonological proximity. It could be responsible in part for the error rate in the unviable context (across the experiments from 6% to 18%).

Although our results make clear that a context-sensitive phonological knowledge of processes is at work, they leave open the question of whether such a mechanism operates at a strictly sub-lexical level (i.e., before lexical access) or whether it is implemented as a more sophisticated, context-sensitive version of a lexical compensation mechanism. Further research involving nonwords will be needed to answer that question.

References

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology, 5*, 183 – 207.

Bailey, P.J., & Summerfield, Q. (1980). Information in speech: Some observations on the perception of [s]-stop clusters. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 536 – 563.

Barry, M. (1992). Palatalisation, assimilation and gestural weakening in connected speech. *Speech Communication, 11,* 393 – 400.

Best, C.T., Morrongiello, B., & Robson, R. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics, 29*, 191 – 211.

Casagrande, J. (1984). *The sound system of French*. Washington D.C.: Georgetown University Press.

Church, K. (1987). Phonological parsing and lexical retrieval. *Cognition*, *25*, 53 – 69.

Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Variation and assimilation in German: consequences for lexical access and representation. *Language and Cognitive Processes*, *16,* 535 – 564.

Content, A., Mousty, P., & Radeau, M. (1990). BRULEX. Une base de données lexicales informatisée pour le français écrit et parlé. *L'Année Psychologique*, 90, 551-566.

Cutler, A., & Norris, D. (1979). Monitoring sentence comprehension. In W.E. Cooper & E.C.T. Walker (Eds.), Sentence Processing : Psycholinguistic Studies Presented to Merrill Garrett. Hillsdale (N.J.) : Erlbaum.

Darcy, I., Peperkamp, S., & Dupoux, E. (in press). Bilinguals play by the rules. Perceptual compensation for assimilation in late L2- learners. *Papers in Laboratory Phonology, 9*.

Dell, F. (1995). Consonant clusters and phonological syllables in French. *Lingua, 95*, 5 – 26.

Ellis, L., & Hardcastle, W.J. (2002). Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics, 30*, 373 – 396.

Féry, C. (2003). Gradedness in the segmental correlates of focus and prosodic phrasing in French. In T. Meisenburg, & M. Selig (Eds.), *Nouveaux départs en phonologie*. Tübingen: Narr.

Fowler, C.A. (1996). Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America, 99*, 1730 – 1741.

Fowler, C.A., & Brown, J.M. (2000). Perceptual parsing of acoustic consequences of velum lowering from information for vowels. *Perception & Psychophysics, 62*, 21 – 32.

Gaskell, G. (2003). Modelling regressive and progressive effects of assimilation in speech perception. *Journal of Phonetics, 31*, 447 – 463.

Gaskell, G., Hare, M., & Marslen-Wilson, W.D. (1995). A connectionist model of phonological representation in speech perception. *Cognitive Science, 19*, 407 – 439.

Gaskell, G., & Marslen-Wilson, W.D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 144 – 158.

Gaskell, G., & Marslen-Wilson, W.D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 380 – 396.

Gaskell, G., & Marslen-Wilson, W.D. (2001). Lexical ambiguity resolution and spoken word recognition: bridging the gap. *Journal of Memory and Language, 44,* 325 – 349.

Gow, D.W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language, 45*, 133 – 159.

Gow, D.W. (2002a). Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance, 28*, 163 – 179.

Gow, D.W. (2002b). Integrating feature cues over time to resolve ambiguity. In S. Hawkins & N. Nguyen (Eds). *Temporal Integration in the perception of speech (TIPS)*. ISCA Tutorial & Research Workshop (ITRW), Aix-en-Provence, France, April 8-10, 2002, ISCA Archive, paper P1-10. http://www.isca-speech.org/archive/tips_02

Gow, D.W. (2003). Feature Parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics, 65*, 575 – 590.

Gow, D.W., & Gordon, P.C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 344 – 359.

Gow, D.W., & Im, A. M. (2004). A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language, 51*, 279 – 296.

Gow, D.W., & Zoll, C. (2002). The Role of Feature Parsing in Speech Processing and Phonology. In Anikó Csirmaz, Zhiqiang Li, Andrew Nevins, Olga Vaysman and Michael Wagner, eds., *Phonological Answers (and their corresponding questions)*. MITWPL 42, MIT Working Papers in Linguistics, Cambridge MA, pp. 55-68.

Grainger, J., & Jacobs, A.M. (1996). Orthographic processing in visual word recognition : a multiple read-out model. *Psychological review, 103,* 518 – 565.

Gumnior, H.. Zwitserlood, P. & Bölte, J. (2005) Assimilation in existing and novel German compounds. *Language and Cognitive Processes* 20, 465 – 488.

Hallé, P., Chereau, C. & Segui, J. (2000). Where is the /b/ in “absurde” [absyrd]? It is in French listeners’ minds. *Journal of Memory and Language*, *43*, 618 – 639.

Hallé, P., Segui, J., Frauenfelder, U., & Meunier, C. (1998). The processing of illegal consonant clusters: a case of perceptual assimilation ? *Journal of experimental psychology: Human perception and performance, 24,* 592 – 608.

Hayes, B. (1984). The phonetics and phonology of Russian voicing assimilation. In M. Aronoff, & R.T. Oehrle (Eds.), *Language sound structure* (pp. 318 – 328). Cambridge (MA): MIT Press.

Hodgson, P., & Miller, J.L. (1996). Internal structure of phonetic categories: Evidence for within-category trading relations. *Journal of the Acoustical Society of America, 100*, 565 – 576.

Jakobson, R. (1956). Die Verteilung der stimmhaften und stimmlosen Geräuschlaute im Russischen. In M. Woltner, & H. Bräuer (Eds.), *Festschrift für Max Vasmer*. Wiesbaden: Harassowitz.

Jun, Jongho (2004) Place assimilation. In B. Hayes, R. Kirchner and D. Steriade (eds.) *Phonetically Based Phonology* (pp. 58-86). Cambridge University Press.

Kingston, J. & R. Diehl (1994) Phonetic knowledge. *Language* 70, 419 – 454.

Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonology Yearbook, 2,* 85 – 138.

Koster, C.J. (1987). *Word recognition in foreign and native language*. Dordrecht, The Netherlands : Foris

Kohler, K. (1981) Contrastive Phonology and the acquisition of phonetic skills. *Phonetica* 38, 213 – 226.

Kohler, K. (1984) Phonetic features in phonology: the feature fortis-lenis. *Phonetica* 41, 150 – 174.

Lahiri, A., & Marslen-Wilson, W.D. (1991). The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition, 38,* 245 – 294.

Lahiri, A., & Marslen-Wilson, W.D. (1992). Lexical processing and phonological representation. In D.R. Ladd, & G.J. Docherty (Eds.), *Papers in Laboratory Phonology, 2* (pp. 229 – 260). Cambridge: Cambridge University Press.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven, & N. Warner (Eds.), *Papers in Laboratory Phonology, 7* (pp. 637 – 675). Berlin: Mouton DeGruyter.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431 – 461.

Lotto, A.J., Kluender, K.R., & Holt, L.L. (1997). Perceptual compensation for coarticulation by Japanese quail (Coturnix coturnix japonica). *Journal of the Acoustical Society of America*, *102*, 1134 – 1140.

Mack, M. (1982) Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *The Journal of the Acoustical Society of America 71*, 173 – 178.

Marslen-Wilson, W.D., Nix, A., & Gaskell, G. (1995). Phonological variation in lexical access: abstractness, inference and English place assimilation. *Language and Cognitive Processes, 10*, 285 – 308.

Marslen-Wilson, W.D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*, 29 – 63.

McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1 – 86.

Mehler, J., Sebastian, N., Altmann, G., Dupoux, E., Christophe, A., & Pallier, C. (1993) Understanding compressed sentences: the role of rhythm and meaning. *Annals of the New York Academy of Sciences* 682, 272 – 282.

Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception & Psychophysics, 65*, 956 – 969.

Mitterer, H., Csépe, V., & Blomert, L. (2003). Compensation for phonological assimilation in perception: Evidence from Hungarian liquid assimilation. In M.J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2321 – 2324). Barcelona, 3-9 August 2003.

Nolan, F. (1992). The descriptive role of segments: evidence from assimilation. In D.R. Ladd, & G.J. Docherty (Eds.), *Papers in Laboratory Phonology, 2* (pp. 261 – 280). Cambridge: Cambridge University Press.

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition, 52*, 189 – 234.

Otake, T., Yoneyama, K., Cutler, A., & van der Lugt, A. (1996). The representation of Japanese moraic nasals. *Journal of the Acoustical Society of America, 100,* 3831 – 3842.

Pallier, C., Dupoux, E., & Jeannin, X. (1997). “EXPE: an expandable programming language for on-line psychological experiments.” *Behavior Research Methods, Instruments, & Computers*, *29*, 322-327.

Parker, E.M., Diehl, R.C., & Kluender, K.R. (1986). Trading relations in speech and nonspeech. *Perception & Psychophysics, 39,* 129 – 142.

Pulleyblank, D. (1988). Underspecification, the feature hierarchy and Tiv vowels. *Phonology, 5*, 299 – 326.

Quené, H., van Rossum, M., & van Wijck, M. (1998). Assimilation and anticipation in word perception. In: *Proceedings of the Fifth International Conference on Spoken Language Processing*, Sydney, December 1998. volume 3, pp. 871-874.

Repp, B.H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin, 92,* 81 – 110.

Rigault, A. (1967). L’assimilation consonantique de sonorité en français : étude acoustique et perceptuelle. In B. Hála, M. Romportl, & P. Janotà (Eds.), *Proceedings of the 6th International Congress of Phonetic Sciences,* (pp. 763 – 766). Prague : Academia.

Roca, I., & Johnson, W. (1999). *A Course in Phonology*. Oxford: Blackwell.

Samuel, A.G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General, 110,* 474 – 494.

Samuel, A.G. (1996). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General, 125*, 28 – 51.

Samuel, A.G. (2001). Knowing a word affects the fundamental perception of the sounds within it. *Psychological Science*, *12*, 348 – 351.

Sebastian-Gallés, N., Dupoux, E., Costa, A., & Mehler, J. (2000) Adaptation to time-compressed speech: phonological determinants. *Perception & Psychophysics,* 62, 834 – 842.

Sinnott, J.M., & Saporita, T.A. (2000). Differences in American English, Spanish and monkey perception of the say-stay trading relation. *Perception & Psychophysics*, *62,* 1312 – 1319.

Snoeren, N. D., & Segui, J. (2003). A voice for the voiceless : voice assimilation in French. In M.J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2325 – 2328). Barcelona, 3-9 August 2003.

Snoeren, N, Hallé, P. & Segui, J. (2006). A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics* 34, 241-268.

Stevens, K. (1998). *Acoustic phonetics*. Cambridge (MA): MIT Press.

Summerfield, Q., & Haggard, M. (1977). On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *Journal of the Acoustical Society of America, 62*, 435 – 448.

Treiman, M. (1999). There are two types of psychometric functions: A theory of cue combination in the processing of complex stimuli with implications for categorical perception. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 517 – 546.

Trubetzkoy, N.S. (1958). *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht [first edition 1939].

Weber, A. (2001). Help or hindrance : How violation of different assimilation rules affects spoken language processing. *Language and Speech* 44, 95 – 118.

Wells, J.C. (1982). *Accents of English I*. Cambridge: Cambridge University Press.

Wetzels, W. L., & Mascaró, J. (2001). The typology of voicing and devoicing. *Language, 77*, 207 – 244.

Wiese, R. (1996). *The phonology of German*. Oxford: Clarendon Press.

Wilson, M. D. (1988). MRC Psycholinguistic Database: Machine-usable dictionary, version 2.0. *Behavior Research Methods, Instrumentation, and Computers,* 20, 6-10.

Wright, R. (2004) A review of perceptual cues and cue robustness. In B. Hayes, R. Kirchner and D. Steriade (eds.) *Phonetically Based Phonology* (pp. 34-57). Cambridge University Press.

MRC Psycholinguistic Database : http://www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm
(retrieved 11.07.2006, 13h00)

Author's Note

We thank Sharon Peperkamp for useful discussion and feedback, Uta Frith, Jim Morgan, Joe Pater and John Kingston for providing testing facilities, Sarah White for help in recruiting participants, Zsanna Sebestenny and Katherine White for their help with the British English stimuli. As well, we are grateful to the volume editors Caroline Féry and Frank Kügler, for providing helpful comments and discussion.

This work was supported by a Ph.D. fellowship of the Landesgraduiertenförderung Rheinland Pfalz from Mainz University (I.D.), a Franco-German mobility fellowship from the Robert Bosch foundation (I.D.), a Marie Curie fellowship of the European Community programme Quality of Life (QLG1-CT-1999-51305) (F.R.), an APN grant "Les surdités phonologiques: Etudes interlangues", 2000 (E.D. / Sharon Peperkamp), and a grant from CNRS "Acquisition phonologique précoce : algorithmes et simulations", 2002 (E.D. / Sharon Peperkamp).

Appendix

French words used in experiment 1

Target	Gloss	Un-changed form	Changed form	No-change Context	Unviable Context	Viable Context
				Place		
bête	(beast)	[bɛt]	[bɛp]	nuisible 'cumbrous' [nɥizibl]	feroce 'ferocious' [feʁɔs]	poilue 'hairy' [pwaly:]
boîte	(box)	[bwat]	[bwak]	marron 'brown' [maʁɔ̃]	fermee 'closed' [fɛʁme:]	carrée 'square' [kaʁe:]
botte	(boot)	[bɔt]	[bɔp]	montantes 'high' [mɔ̃tɑ̃t]	rayées 'striped' [ʁɛje:]	pointue 'spiky' [pwɛ̃ty:]
chouette	(owl)	[ʃwɛt]	[ʃwɛk]	malade 'sick' [malad]	sauvage 'wild' [sovaʒ]	craintive 'frightened' [kʁɛ̃tiv]
dune	(dune)	[dyn]	[dym]	lointaine 'remote' [lwɛ̃tɛn]	sauvage 'wild' [sovaʒ]	brumeuse'brumous' [bʁymøz]
guide	(guide)	[gid]	[gib]	raciste 'racist' [ʁasist]	vulgaire 'vulgar' [vylgɛʁ]	bourru 'grouchy' [buʁy]
lune	(moon)	[lyn]	[lym]	jaune 'yellow' [ʒon]	rousse 'red' [ʁus]	pâle 'pale' [pɑl]
mode	(fashion)	[mɔd]	[mɔg]	locale 'local' [lokal]	zoulou 'Zulu' [zulu]	guerrière 'combat' [gɛʁjɛʁ]
moine	(monk)	[mwan]	[mwam]	rusé 'wily' [ʁyze]	serviable 'helpful' [sɛʁvjabl]	bavard 'talkative' [bavaʁ]
prune	(plum)	[pʁyn]	[pʁym]	juteuses 'juicy' [ʒytøz]	sucrées 'sweet' [sykʁe:]	pourries 'rotten' [puʁi:]
reine	(queen)	[ʀɛn]	[ʀɛm]	généreuse'generous'[ʒeneʁøz]	respectée'respected'[ʁɛspɛkte:]	paresseuse 'lazy' [paʁɛsøz]
ride	(wrinkle)	[ʀid]	[ʀig]	légère 'light' [leʒɛʁ]	discrète 'discreet' [diskʁɛt]	gracieuse 'graceful' [gʁasjøz]
ruine	(ruin)	[ʀɥin]	[ʀɥim]	romaine 'Latin' [ʁomɛn]	célebre 'famous' [selɛbʁ]	baroque 'baroque' [baʁɔk]
stade	(stadium)	[stad]	[stab]	rénové 'renovated' [ʁenove]	démodé 'outdated' [demode]	bétonné 'concrete' [betone]
trone	(throne)	[tʁon]	[tʁom]	rocheux 'rocky' [ʁoʃø]	royal 'royal' [ʁwajal]	princier 'princely' [pʁɛ̃sje]
zone	(zone)	[zon]	[zom]	rurale 'rural' [ʁyʁal]	fluviale 'riverine' [flyvjal]	portuaire 'harbor' [pɔʁtɥɛʁ]
				Voicing		
badge	(badge)	[badʒ]	[batʃ]	métallique 'metallic' [metalik]	ravissant 'charming' [ʁavisɑ̃]	parfumé'perfumed' [paʁfyme]
cape	(cape)	[kap]	[kab]	longue 'long' [lɔ̃g]	neuve 'new' [nøv]	grise 'grey' [gʁiz]
chèque	(check)	[ʃɛk]	[ʃɛg]	mensuel 'monthly' [mɑ̃sɥɛl]	reçu 'received' [ʁəsy]	volé 'stolen' [vole]
couche	(layer)	[kuʃ]	[kuʒ]	neigeuse 'snow' [neʒøz]	marron 'brown' [maʁɔ̃]	jaunie 'yellowed' [ʒoni:]
coude	(elbow)	[kud]	[kut]	meurtri 'injured' [mœʁtʁi]	raidi 'rigid' [ʁɛdi]	tordu 'twisted' [tɔʁdy]
cuve	(tank)	[kyv]	[kyf]	mobile 'mobile' [mobil]	remplie 'full' [ʁɑ̃pli:]	fendue 'ripped' [fɑ̃dy:]
faute	(error)	[fot]	[fod]	majeure 'major' [maʒœʁ]	légère 'light' [leʒɛʁ]	discrète 'discreet' [diskʁɛt]
globe	(globe)	[glɔb]	[glɔp]	miroitant 'mirroring' [miʁwatɑ̃]	lumineux 'luminous' [lyminø]	pailleté 'sequined' [pajəte]
lac	(lake)	[lak]	[lag]	limpide 'clear' [lɛ̃pid]	nordique 'Nordic' [nɔʁdik]	gelé 'frosted' [ʒəle]
lave	(lava)	[lav]	[laf]	mouvante 'moving' [muvɑ̃t]	rugueuse 'cragged' [ʁygøz]	pateuse 'pasty' [pɑtøz]
nappe	(tablecloth)	[nap]	[nab]	rayée 'striped' [ʁɛje:]	rustique 'rustic' [ʁystik]	brodée 'embroidered' [bʁode:]
neige	(snow)	[nɛʒ]	[nɛʃ]	mouillée 'wet' [muje:]	marron 'brown' [maʁɔ̃]	poudreuse 'powder' [pudʁøz]
nuage	(cloud)	[nɥaʒ]	[nɥaʃ]	rosés 'rosy' [ʁoze]	nacrés 'pearly' [nakʁe]	chargés 'loaded' [ʃaʁʒe]
plaque	(plate)	[plak]	[plag]	noircie 'blacked' [nwarsi:]	rouillée 'rust' [ʁuje:]	brillante 'shiny' [bʁijɑ̃t]
robe	(dress)	[ʀɔb]	[ʀɔp]	rouge 'red' [ʁuʒ]	noire 'black' [nwaʁ]	sale 'dirty' [sal]
route	(road)	[ʀut]	[ʀud]	magnifique'beautiful'[maɲifik]	nationale 'main' [nasjonal]	dangereuse'dangerous'[dɑ̃ʒəʁøz]

American words used in experiment 2

Target	Unchang-ed form	Changed form	No-change context	Unviable context	Viable context
			Place		
bad	[bæd]	[bæb]	[dɪʃ] dish	[lʌnʃ] lunch	[bɪə] beer
fat	[fæt]	[fæp]	[mʌŋki:] monkey	[skwɪɹɪl] squirrel	[pʌpi] puppy
great	[gɹeɪt]	[gɹeɪk]	[faɪt] fight	[mætʃ] match	[kɹu:z] cruise
mad	[mæd]	[mæb]	[mɒðɜ:] mother	[dɒtɜ:] daughter	[bɹɒðɜ:] brother
red	[ɹed]	[ɹeg]	[nekleɪs] necklace	[lɪpstɪk] lipstick	[glæsɪz] glasses
sad	[sæd]	[sæb]	[mʊvi:] movie	[nɒvəl] novel	[bəleɪ] ballet
sweet	[swi:t]	[swi:k]	[tʃʌkleɪt] chocolate	[lɪkjɜ:] liqueur	[kʌkteɪl] cocktail
wet	[wet]	[wep]	[ʃu:z] shoes	[sɒks] socks	[pænts] pants
clean	[kli:n]	[kli:m]	[fɔ:k] fork	[spu:n] spoon	[pæn] pan
fun	[fʌn]	[fʌŋ]	[deɪ] day	[naɪt] night	[geɪm] game
green	[gɹi:n]	[gɹi:ŋ]	[veɪz] vase	[tʃæ:] chair	[kʌp] cup
lean	[li:n]	[li:m]	[laɪn] line	[ʃeɪp] shape	[bæk] back
own	[əʊn]	[əʊm]	[laɪf] life	[tʃɔɪs] choice	[plæn] plan
plain	[pleɪn]	[pleɪŋ]	[tʃæpəls] chapels	[tʃɜ:tʃɪz] churches	[kɒndəʊz] condos
tan	[tæn]	[tæm]	[skɑ:f] scarf	[ʃɜ:t] shirt	[belt] belt
thin	[θɪn]	[θɪm]	[nəʊtbʊk] notebook	[li:flet] leaflet	[pækɪt] packet
			Voicing		
big	[bɪg]	[bɪk]	[laɪthaʊs] lighthouse	[ɹɪvɜ:] river	[faʊntɪn] fountain
blind	[blaɪnd]	[blaɪnt]	[leɪdi:] lady	[lɔ:jɜ:] lawyer	[tʃelɪst] cellist
brave	[bɹeɪv]	[bɹeɪf]	[məri:n] marine	[laɪfgɑ:d] lifeguard	[faɪɜ:mæn] fireman
drab	[dɹæb]	[dɹæp]	[laɪtɪŋ] lighting	[meɪkʌp] make-up	[peɪntɪŋ] painting
good	[gʊd]	[gʊt]	[lʊks] looks	[lʌk] luck	[frendz] friends
huge	[hju:dʒ]	[hju:tʃ]	[mænɜ:] manor	[mænʃən] mansion	[fɒɹest] forest
mild	[maɪld]	[maɪlt]	[naɪts] nights	[ɹeɪn] rain	[spɹɪŋ] spring
wise	[waɪz]	[waɪs]	[li:dɜ:] leader	[ɹaɪtɜ:] writer	[ti:tʃɜ:] teacher
best	[best]	[bezd]	[mu:v] move	[ɹʌn] run	[deɪ] day
black	[blæk]	[blæg]	[ɹʌg] rug	[ɹæg] rag	[glʌv] glove
cheap	[tʃi:p]	[tʃi:b]	[lʌnʃ] lunch	[ɹu:m] room	[dɹɪŋk] drink
flat	[flæt]	[flæd]	[ɹɑ:ft] raft	[ɹʌk] rock	[dæm] dam
French	[frentʃ]	[frendʒ]	[meɪd] maid	[nɜ:s] nurse	[gaɪd] guide
nice	[naɪs]	[naɪz]	[mædəʊz] meadows	[ɹeɪlɪŋz] railings	[gɑ:dənz] gardens
thick	[θɪk]	[θɪg]	[ɹəʊp] rope	[laɪn] line	[bɑ:] bar
tough	[tʌf]	[tʌv]	[lesən] lesson	[ɹɪkwest] request	[dɪmænd] demand