Bilinguals play by the rules: perceptual compensation for assimilation in late L2-learners

I. Darcy[1,2]; S. Peperkamp[1,3], E. Dupoux[1]

[1]Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole des Hautes Etudes en Sciences Sociales, Ecole Normale Supérieure, Centre National de la Recherche Scientifique
[2] Universität Potsdam, Institut für Linguistik
[3] Université Paris 8 Saint Denis

Short title : acquiring compensation patterns in L2
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

to appear in:
J.Cole & J. Hualde (eds.) (2005) Papers in Laboratory Phonology 9

Abstract

Phonological rules introduce variation in word forms, that listeners have to compensate for. We previously showed (Darcy 2002, Darcy et al., *under review*) that compensation for phonological variation in perception is driven by language-specific mechanisms. In particular, English speakers compensate more for place assimilation than for voicing assimilation, whereas the reverse holds for French speakers. English indeed has a rule of place assimilation, whereas French has a rule of voicing assimilation. In the present study, we explore the patterns of compensation for assimilation in English learners of French and in French learners of English. We use the same design and stimuli as Darcy 2002, Darcy et al. (under review); in this design, listeners are engaged in a word detection task on sentences containing occurrences of both place assimilation and voicing assimilation. We test British English and American English learners of French as well as French learners of American English on both their native language (L1) and their second language (L2). The results show that beginners interpret their L2 in exactly the same way as their L1: they apply the native compensation pattern to both languages. Advanced learners, by contrast, succeed in compensating for the non-native assimilation rule in their L2, while keeping the native compensation pattern for L1; as little or no interference from L2 on L1 is observed for these learners, we conclude that two separate systems of compensation for phonological processes can co-exist.

1. Introduction

The word recognition system copes easily with the great variability of spoken language. One source for this variability comes from the mutual influence of sounds at word edges, resulting in assimilation: one sound takes over some properties of the neighbor sound. Such cases obscure the direct relationship between a word's surface form and its identity, making word recognition harder. In order to recover from the change and activate the right word representation, assimilation has to be compensated for by the recognition system. This paper addresses the issue of whether late learners compensate for an assimilation rule that is absent from their native language.

Language-specific compensation for phonological processes

The adjustments of phonological structure made by assimilation processes are specific to a particular language. For instance, French has a tendency towards regressive voicing assimilation, whereas English has a tendency towards regressive assimilation of place of articulation. In French, a word like *botte* [bot] 'boot' is produced as [bod] when followed by *grise* [griz] 'grey' (*bo[**dg**]rise*), but not when followed by *mauve* [mov] 'purple' (*bo[**tm**]auve*). French voicing assimilation applies to obstruent clusters inside the same phonological phrase, and propagates the voicing feature (voiced or unvoiced) of the second segment onto the first one. The initial m of *mauve* is not an obstruent, hence it does not trigger assimilation on the last consonant of *botte*. Similarly, in English, coronals (t, d and n) adjust their place of articulation to that of any following labial or velar segment. For example, the word *sweet*, if followed by *melon*, may assimilate the labial place of articulation, and be pronounced as *swee[**pm**]elon*. A word like *grapes* would equally alter the final coronal of *sweet*, which would become velar (*swee[**kg**]rapes*). Such modifications occur at word boundaries: the form of a word may vary depending on the words surrounding it. Consequently, they are a problem for word recognition, and have to be learned specifically for each language.

Listeners have been shown to compensate for the assimilation processes that occur in their native language in a very precise way (Gaskell & Marslen-Wilson, 1996, 1998): When presented with a changed word (*freigh[p]*), listeners were able to detect the target phoneme /t/ only when that changed word was followed by a context legitimating the change as assimilation (for example, in the phrase *freight bearer* : *freigh[pb]earer*). They failed to detect /t/ in the phrase *freight carrier* where the change is rendered inappropriate (unviable) by the context (*freigh[pk]arier*). Here, the change does not correspond to any existing rule in English, and is not compensated for by native listeners of English. Evidence for language-specific processing of assimilated sounds has been obtained using different methods, languages and processes, and mainly through presenting non-native input or illegal

sequences to participants (Coenen, Zwitserlood & Bölte, 2001; Koster, 1987; Mitterer & Blomert, 2003; Otake, Yoneyama, Cutler, & van der Lugt, 1996; Weber, 2001). One study (Darcy et al., under review) investigated compensation for assimilation avoiding the problem of non-native of illegal input, and showed that listeners compensate only for those changes which correspond to existing processes in the language, and not for others: Using a word detection task, British, American and French listeners were presented – exclusively in their native language – to different kinds of alteration on target words, corresponding or not to an existent assimilation process in that language: Sentences *for each language* provided examples of both the English place assimilation process and of the French voicing assimilation process. Changes were complete rather than gradient[1]. Listeners compensated more for the native process than for the non-native one. French subjects, for instance, detected the target [bɔt] more often in *bo[dg]rise* than in *bo[dm]auve*, compensating for voicing assimilation when appropriate, whereas they did respond less to the target *lune* [lyn] 'moon' in both conditions (appropriate) *lu[mp]ale* and (inappropriate) *lu[mʁ]ousse*, hence correctly rejecting a place assimilated target, independently of the context condition. Similarly, both British English and American English listeners detected appropriately assimilated targets only in the native case, i.e. for place assimilation. Their detection of voiced assimilated targets was significantly less important for both the appropriate and the inappropriate context.

One of the most striking conclusions of the Darcy et al. study is that compensation for assimilation reflects language-specific phonological knowledge. This raises the question to what extent adult learners are able, when they are faced with a second language and a different phonological system, to learn these phonological competences for purposes of processing L2, and in case they do learn, to what extent does this learning influence their L1-phonological knowledge.

Acquisition of phonological processes

We briefly introduce the two main views on late learners' acquisition of phonology: on one hand, a large body of evidence in second language learning suggests that phonological properties of the target language are very difficult to acquire, even in learners who were exposed to the new system early in

[1] Although spontaneous realizations of assimilation are most of the time gradient and cannot be described as a simple phoneme substitution (Nolan, 1992; Ellis & Hardcastle, 2002), this does not imply that speech perception is also gradient: listeners have great difficulties in accessing fine-grained phonetic cues, unless these cues yield a change in phoneme category. The real issue, therefore, is how often assimilated tokens cross a perceptual category boundary. If even a small proportion does, this creates a compensation problem no different from the one that would have been created by the presence of discrete deliberate substitutions. Several studies have shown that the perceptual apparatus displays similar robust and numerically large compensation for natural assimilations and deliberate substitutions, even in cases where the underlying process in production has been shown to be gradient (Nolan, 1992; Gaskell & Marslen-Wilson, 1996, 1998).

life. Flege, Yeni-Komshian and Liu (1999) for example have shown that listeners were able to detect a slight foreign accent in Korean learners who arrived in the USA as early as 1 year of age. Similarly, Pallier et al. (1997) showed that Catalan-Spanish bilinguals who were born in a monolingual family but immersed in a bilingual culture from age 4 still have problems in perceiving aspects of their non-native phonologies.

This difficulty is either attributed to maturational factors (critical period hypothesis, applying to other linguistic domains also, Long, 1990; Flege, 1995; Flege et al., 1999; Johnson & Newport 1989; Weber-Fox & Neville, 1996), or to interference/competition effects between L1 and L2 (Flege, Frieda & Nozawa, 1997): When faced with a non-native phonological system, the non-native listener would process the input using L1's phonological system, thereby showing intrusion of L1 into L2-processing. This is among others personified as the Perceptual Assimilation Model (Best, 1995), i.e. perceptual assimilation of the non-native phonetic categories to those of L1 (Kuhl, 1991; Flege, 1995) and has been shown in a number of different studies (e.g. McAllister, Flege & Piske, 2002; Strange et al., 1998; Hallé, Best & Levitt, 1999).

On the other hand, experience in a second language (as measured for example by proficiency) may modify native speech perception (Flege, Bohn & Jang, 1997; Weber, 2001), and even short exposure to ambiguous sounds in a specific context may provoke a category boundary shift (Norris, McQueen & Cutler, 2003). Examples such as these could be interpreted as evidence for the malleability of phonological knowledge, which allows for modifications through L2-exposure. In some cases, even native attainment in phonology has been reported, based mainly on observations made in the domain of foreign accent judgments in production of non-native speakers (Bongaerts, 1999; Flege et al., 1999).

The present study

So far, little work has been done specifically on the acquisition of phonological processes. What would native speakers of English learning French do when they process French sentences? We use the same design and stimuli as in Darcy et al. In Experiment 1, we report the results of two groups of British English native speakers hearing French sentences. The first group consists of beginning learners of French, whereas the second is more experienced. In Experiment 2, we extend the exploration of learners' compensation patterns for assimilation using a crossover design. We test American learners of French and French learners of American English on both their first and second language. This series of experiments allows us to address how learners represent the phonological system of L2. This is an important issue for better understanding how words are recognized in a foreign language, and how different languages are represented in a learner's brain.

Three possibilities arise: First, participants might apply the same behavior as in L1 (more compensation for the native process over the non-native one) when hearing L2 and compensate in both languages for changes native in L1. In this case, the phonological system is not modifiable and proficiency in L2 does not play any role.

Second, processes applying in L2 are learnable, but if this learning is only extending the L1-phonological system, the prediction is that more advanced learners would apply compensation to both the native and non-native process in L1 and L2.

These two possibilities would not prevent the overapplication of L1-rules in L2-processing. In the second case, the newly acquired L2-processes would also alter the first-language-system, in the same way as phonetic categories for L1 and L2 sharing the same space influence each other.

It is only when this learning is combined with the possibility to build separate systems that a native-like compensation pattern would be possible in both tested languages. In this case, when hearing L2, highly proficient learners would detect the correctly assimilated tokens, and correctly reject the inappropriately assimilated ones according to L2-phonology, thus showing the same pattern as native listeners of that language.

2. Experiment 1

Method

Stimuli

The stimuli used in all experiments presented in this paper are those described in Darcy et al. (under review). The French and the American English stimuli were identical in design and conditions. We describe here the construction of French stimuli only: Thirty-two target items were selected. They are all French monosyllabic nouns, with a C(C)VC structure. Each target (e.g. “robe”) is modified on its last consonant according to voicing or place feature and ends up with 2 forms, the “original” and the “changed” form, which is always a non-word (nw), e.g. /rob/ vs. /rop/$_{nw}$. The target in either form is then paired to three different context words, producing three different conditions. Table 1 illustrates these conditions with two examples, for voicing “robe” (‘dress’), and for place “lune” (‘moon’).

Table 1: overview of context conditions and changes for French stimuli. Here, voicing assimilation is the native process, whereas place assimilation is non-native. For English stimuli, this pattern is reversed.

type of change / condition	target "robe" [ʀɔb] 'dress'			target "lune" [lyn] 'moon'		
	Native (voicing)			Non-native (place)		
appropriate change (viable context)	1a	ro[**p**+s]ale	'dirty dress'	2a	lu[**m**+p]ale	'pale moon'
non appropriate change (unviable context)	1b	ro[**p**+n]oire	'black dress'	2b	lu[**m**+ʀ]ousse	'red moon'
no-change	1c	ro[b+r]ouge	'red dress'	2c	lu[n+ʒ]aune	'yellow moon'

The changed form (nw) is associated with 2 context words whose initial consonant either does trigger assimilation (e.g. 1a and 2a, for voicing or place respectively), or does *not* trigger them (e.g. 1b and 2b). Trigger contexts (a) for voicing are obstruents agreeing in voicing with the changed form's final consonant, for place, they agree in place (labial or velar). The unchanged target (baseline c: *no-change* condition) is paired to a third context word and does not undergo assimilation (1c and 2c). Conditions a and b share the same target form (changed form), but differ in the context. If the context is a trigger, the change is appropriate (a: *viable context* condition); if it is not, the change is consequently not appropriate (b: *unviable context* condition). Conditions b and c on the other hand share the same kind of context, but differ in the form of the target (changed vs. non changed).

The whole stimuli set consists of two sets of 16 frequency-matched target items: the Voicing Set and the Place Set. In the Voicing Set, all 16 targets end in a final obstruent (voiced for half of the items and unvoiced for the other half). Their corresponding changed forms were constructed by switching the voicing feature of the final obstruents (e.g. *robe* /rob/ 'dress' - *rope* /rop/$_{[nw]}$, or *lac* /lak/ 'lake' - *lague* /lag/$_{[nw]}$). In the Place Set, all final consonants are coronal; half are nasals and half are stops. Sixteen matched non-words were obtained by a change in the place feature of the final consonant (e.g. *lune* /lyn/ 'moon' - *lume* /lym/$_{[nw]}$ or *boite* /bwat/ 'box' - *boike* /bwak/$_{[nw]}$).

Next, the three pairs "target form – context word" associated to each target item are embedded in three sentence frames (i.e. a sentence beginning and ending), where each of the three pairs may be inserted and results in a plausible sentence (e.g. for the item "robe", one of the three sentences is *Elle a mis sa* ___ ___ *aujourd'hui* 'she put on her ___ ___ today', where the three pairs 1a, 1b and 1c from Table 1 are inserted). Globally, the sentence frames are matched in number of words and position of the insertion slots across the Voicing Set and the Place Set. Combining pairs and frames gives rise to nine sentences associated with each item. We thus obtain 32 targets x 3 pairs x 3 frames = 288 actual sentences in total. All sentences are fully counterbalanced across three experimental lists. Thirty additional training and filler sentences were constructed, which do not include any case of viable or unviable assimilation (voicing or place).

Changes were rendered orthographically in the printed lists of sentences (e.g. for /rop/$_{[nw]}$, the printed sentence was "elle a mis sa **rope** sale") used for recordings. For each language, a female native speakers was trained to pronounce the changed form of targets in a similarly natural way as the unchanged form (producing minimal pairs differing on the last phoneme), until familiar with all sentences. They were instructed to pronounce sentences in a fast speech rate, with maximal naturalness, and without emphasis on changed segments or on the preceding vowels. Within the critical sequences C+C [target's final consonant + context-word's initial], the first consonant was never released orally. Intonation was kept the same for all sentences of each frame. All sentences were recorded as a whole, several times. Finally, the set of 9 sentences per item which were most natural and where changes were clearest was manually selected according to ratings performed by two trained phoneticians. Single target words were recorded by a male native speaker.

Separate experiments (forced choice task on the last consonant of the target items in their changed or unchanged forms without context, see Darcy et al.) ensured that in isolation, all final consonants were perceived as the intended ones regardless of whether they were changed or unchanged and whether the change had been produced in a viable or unviable context.

Participants

None of the participants tested in this and the following experiments had previously taken part in a similar experiment, nor reported any history of hearing disorders. None of them grew up bilingually. In total, 46 native English speakers were tested (in Paris or London) in experiment 1 and paid for participation. The 19 participants living in Paris were tested only on French sentences. However, the 27 Londoners were tested on both British English (reported in Darcy et al.) and French sentences (in a single session). The London-group is homogenous: All were native speakers of south-eastern British English. The Parisian group of English listeners is less homogenous. Only four of them were raised in the south-east of England. The others came from different origins. Both groups are similar in terms of first exposure to French (10,5 vs. 11,1 years) and length of instruction at school (on average 5 vs. 6,5 years). They however differ with respect to French-exposure: None of the Londoner have had extensive exposure to French nor spent time living in France, whereas the Parisian group was living in France for at least 12 months (mean length of residence: 7,4 yrs.). We label the London-group as "beginning learners", and the Paris-group as "advanced learners". All participants reported having understood every sentence after the experiment.

Procedure

A test trial consists of a target word presented in isolation (male voice), followed after 500 ms of silence by a sentence (female voice). For the group tested in London (tested on both English and French), we used the same method for both languages: Participants have to press a "yes" or a "no" button as a response to the question "Do you think that the word presented alone is the *same* in the following sentence?" This instruction – together with a specific training – was given in order to draw their attention to the *goodness of pronunciation* of words in the sentence, i.e. to the *form* of words. For the same reason only a few filler words were included. The test phase was split into three blocks of 36 randomized trials (a given test item appeared only once within each block). The Parisian group was tested with the same method, except that instead of the "yes/no" response type used for the beginning learners, participants here had to press "yes" or refrain from pressing a button. This is the method used for French subjects hearing the same French sentences, in Experiment 1 of the Darcy et al. study.

The experiment lasted 20 minutes. Participants were tested individually. Our main measure is the compensation index (formula 1) for each subject and item, computed on the basis of the number of yes-responses as a function of condition and contrast type (place vs. voicing).

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

This index calculates the relative value of detection in the viable condition as a function of both other conditions. This allows us to compute 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.

Results

We present the results for each proficiency-group. For these and all the following results, the same criterion for item and participant exclusion as in Darcy et al. applies: items causing more than 50% false alarms or misses in the native speakers results on L1 are dropped. For French stimuli, one voicing item is excluded. The error rate for each participant is based on the performance on the no-change and unviable change conditions alone. Participants who made more than 37.5% errors for either the Voicing or the Place items were considered as failing to perform the word detection task for this contrast and replaced. Four beginning and five advanced learners were excluded. We computed

the compensation index for each participant and each item. Compensation indices are displayed in Figure 1, for the group of beginning learners hearing English (ALL, N=27) and French (N=23) sentences, as well as for the group of advanced learners hearing French (N=14). For this and the following experiments, detection rates in place or voicing for each condition are summarized in the Appendix.

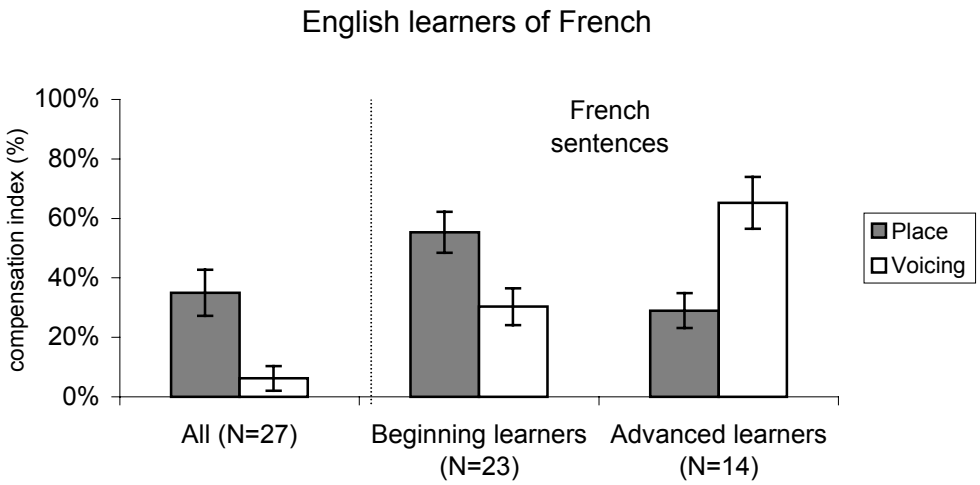

Fig. 1. Compensation indices (%) for English listeners, hearing English sentences (N=27) and French sentences (beginning learners, N=23, and advanced learners, N=14)

For the beginner group, mean index is 55% for place and 30% for voicing. Index means were used as the dependent variable in an ANOVA with Contrast (place vs. voicing) as a within-subject (respectively between-items) factor. The difference is significant by subjects, not by items because of high variability (standard error 0,16) on place items ($F_1[1,22]=11.4$, $p<.003$; $F_2[1,29]=1.8$, $p>.1$). This effect shows that all beginning learners, even when hearing French, compensate more for place assimilation than for voicing.

Advanced learners have a mean compensation index of 65% for voicing, and 28% for place. This difference is significant by subjects and items ($F_1[1,13]=12.1$, $p<.004$; $F_2[1,29]=17.3$, $p<.0001$), indicating that this group compensates more for voicing than place when hearing French.

On the same sentences, both groups behave in a different way, yet their pattern of compensation is reversed. An ANOVA on indices including both groups and declaring the factors proficiency (beginning vs. advanced) and contrast (place vs. voicing) revealed a significant interaction by subjects and by items ($F_1[1,35]=24.1$, $p<.0001$; $F_2[1,29]=12.9$, $p<.001$). This comparison shows that proficiency may be responsible for the different response to place vs. voicing in both groups.

Discussion

In this experiment, two groups of subjects tested on the same sentences showed a different pattern of compensation according to proficiency in a second language. The beginner group compensates more for a native process in their phonology, than for a non-native one, independently of the language heard. They apply L1-phonology in L2-processing. By contrast, more advanced learners of French compensate more for the process which is native in L2, and less for a process they know from their first language. They behave like the French native listeners who hear French. These results seem to contradict the first hypothesis: compensation *is* adaptable to L2-processes, and this adaptation could be correlated to proficiency in that second language. In fact, correlation analysis on compensation indices according to the length of residence in France for the advanced group was significant ($r=0.6$, $F[1,12]=7.1$, $p=.019$, higher index for voicing correlates to a longer stay in France).

The combined results of experiment 1 suggest that beginners tend to interpret assimilation alternations according to L1-phonology, whereas more advanced learners do this according to L2-phonology, thus having a native-like compensation behavior in L2 (i.e. similar to native listeners of that L2). However, for the advanced learners, there is no way to tell what the compensation pattern would look like in L1. Before we can determine whether there are two or only one system, according to our predictions above, we would have to directly compare the pattern of results of this proficient group on both French and English sentences. This was difficult due to the mixed origins of participants and differences in testing methods. Experiment 2 was designed in order to compare compensation within the same subjects on both languages, and to investigate the effect of proficiency in L2 on compensation pattern in each language.

3. Experiment 2

Method

Procedure

The procedure used here is the same as in Experiment 1 above (for beginning learners): Participants are requested to give a "yes/no" response. Before starting, each participant filled in a language-background questionnaire, and was assigned to a group and language-order. Instructions were given before each part in the tested language. The whole procedure lasted about 1 hour.

Participants

In total, 58 subjects were tested in their country of residence (either France or the United States), 26 Americans (19 in Paris, 7 in the U.S.) and 32 French (10 in Paris, 22 in the U.S.). They were paid for participation. Of the 10 French tested in Paris, 2 were tested on French sentences only, and 8 on American English sentences only. These subjects are excluded from "within-subject"-comparison analyses. In total, 24 French were tested on French, and 30 were tested on American. All 22 French tested in the U.S. and all American participants were tested on both languages.

Participant's biographical data include length of residence in L2-speaking country, amount of interaction in L2 for different situations, the importance of L2 in life, and age of acquisition of L2. Table 2 summarizes for both groups the mean values of each factor.

Table 2: summary of main biographical characteristics for Americans and French participants

	N	Mean age	*std. dev.*	Age of 1st expo. to L2	*std. dev.*	Length of residence (yrs)	*std. dev.*	Amount of interaction in L2 (%)	*std. dev.*	Importance of L2 in life (%)	*std. dev.*
French	32	26,1	6,4	9,7	3,7	3,8	5	48%	0,3	28%	0,17
Americans	26	28,9	11,0	12,5	4,4	3,7	7	34%	0,3	28%	0,18
Difference		F(1,56)=1.19, p>.1		F(1,56)=2.56, p<.01		F(1,56)=0.19, p>.8		F(1,56)=-1.5, p>.1		F(1,56)=-0.01, p>.9	

Proficiency measure

We are interested in assessing some sort of phonological proficiency, which is possibly related to the amount of L2-native input. Indeed, perceptual learning of this kind of phonological competence could reasonably be initiated through repeated exposure to occurrences of assimilation, that is, through intensive exposure to L2-native input.

As an objective measure of proficiency we used Length of Residence in L2-country (LoR), in months. We separated both groups in long LoR vs. short LoR on the basis of an arbitrary threshold: For Americans, long LoR is more than 24 months residence in France, short LoR being less than 24 months. For French, due to numerical balance reasons, long LoR was longer than 36 months as a resident in the United States. Table 3 shows the mean LoR-values for the different groups. No significant difference can be seen between both populations (American vs. French) within the same LoR-range, indicating that the different separation threshold (24 vs. 36 months) is not confounding.

Table 3: difference between both groups for comparable Length of Residence

Mean	Long LoR	range	N	Short LoR	range	N
Americans (> 24 months)	116,3	30-396	9	6,7	0-22	17
French (> 36 months)	106,6	48-180	10	11,3	0-36	22
Difference	n.s.			n.s.		

Correlation analyses showed that neither country of residence nor participant's age correlated to residence length. In contrast, both the amount of interaction in L2 and the global importance of L2 in life did correlate to the LoR. Thus if high L2-proficiency is induced by intensive interaction in L2 in daily life, using LoR as a measure of proficiency seems to be appropriate. This separation of both groups resulted in the labels "high proficient" or "advanced learners" for long LoRs, and "less proficient" or "beginning learners" for short LoRs. The label "less proficient", however, should not be misleading – all participants were very good at speaking and understanding L2, and could perfectly understand every L2-sentence presented in the experiments.

In the following section, results are presented first for L1 in both groups, then for L2, split according to proficiency. The results of Americans hearing American English sentences are those reported in Experiment 3 of the Darcy et al. study.

3.2. Results and Discussion

3.2.1. American hearing American English

Due to space limitations, we will summarize here only the most crucial results. Four items were rejected, no participant was excluded from the analyses. As in previous experiments, we calculated the compensation index for each participant and each contrast (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 vs. voicing). The effect of contrast is significant by participants, not by items (see discussion below), confirming that all subjects behave similarly and compensate significantly more for place than voicing assimilation ($F_1[1,25]=57$, $p<.0001$; $F_2[1,26]=2.7$, $p>.1$) when hearing American English. Compensation indices are displayed in Figure 2. The results on American English sentences for the whole group of American listeners is displayed in the left part of the graph.

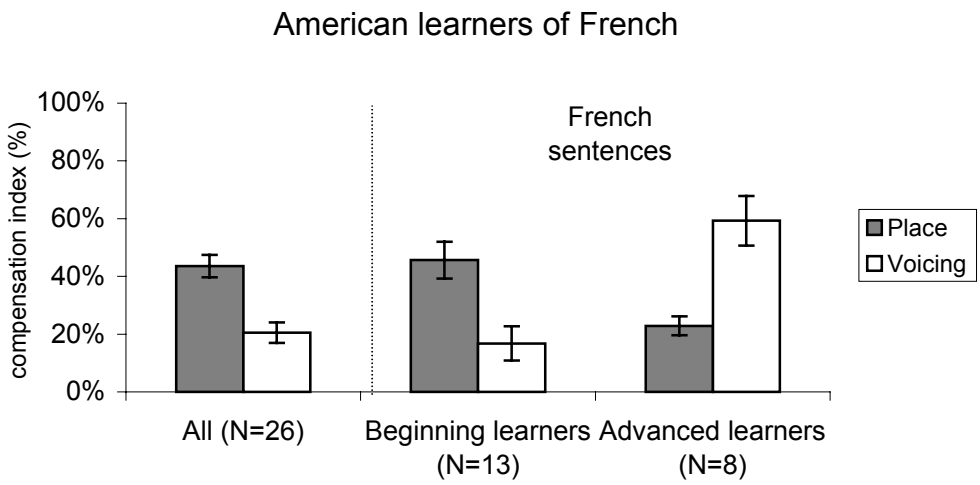

Fig. 2. Compensation indices (%) for American listeners, hearing American English sentences (N=26) and French sentences (beginning learners, N=13, and advanced learners, N=8).

This group presents a pattern of results similar to that of French and British native listeners (Darcy et al.). Crucially, the degree of compensation by subjects is higher for a place assimilation process than for the non-native voicing assimilation occurrences. The lack of significant effects for items can be explained by the fact that some items did elicit more compensation than others. Crucially, items in the voicing set were behaving differently depending on their status as "voicing" or "devoicing" item. Indeed, devoicing items (*big* in *bi[kf]ountain*) were largely compensated, whereas voicing items (*flat* in *fla[dd]am*) were not. This could reflect compensation for a process of partial phonetic final devoicing applying in American English (Hyman 1975, Keating 1984). 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 ($F_1[1,25]=34.5$, $p<.0001$; $F_2[1,19]=8.8$, $p<.008$).

The results for this group serves as comparison basis for evaluating the performance of the same subjects hearing French sentences presented below. We first consider the results of the French group hearing French.

3.2.2. French hearing French

One voicing item and one participant were excluded. Mean compensation index is 32% for the place type of change, and of 70% for the voicing contrast. Although compensation is not null for a non-native kind of change (32% for place), it is significantly less important than for the native process ($F_1[1,22]=32.9$, $p<.0001$; $F_2[1,29]=18.8$, $p<.0001$).

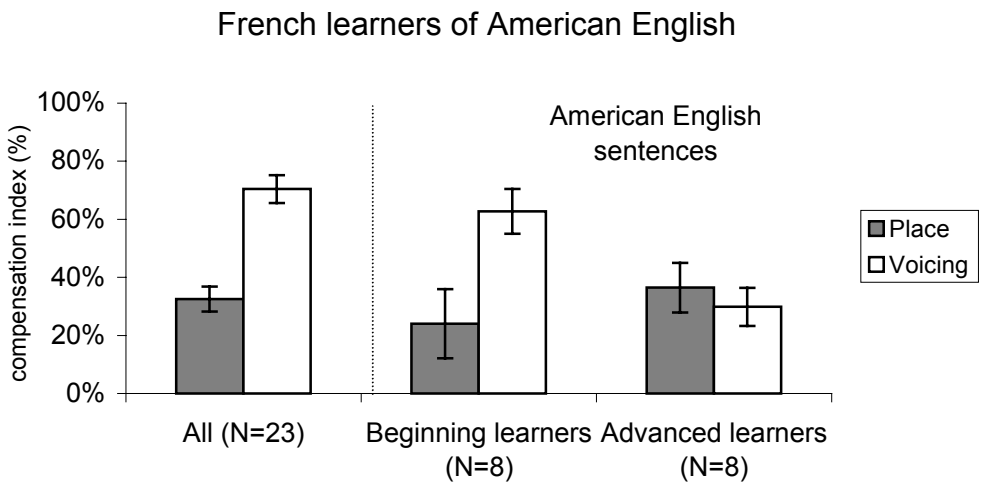

Fig. 3. Compensation indices (%) for French listeners, hearing French sentences (N=23) and American English sentences (beginning learners, N=8, and advanced learners, N=8).

Fig. 3 displays the compensation indices for both French and American English sentences. Compensation indices for French sentences are displayed in the left part of the graph. These results replicates the previously obtained pattern with same sentences but another French group and a slightly different procedure, reported in Darcy et al. Another difference to the first group is that these French listeners are also late learners of English. However, this shouldn't radically influence their processing of French if there are two systems. Furthermore, the present results are similar to the pattern obtained for the American group hearing American English: both groups compensate more for a native process than for a non-native one, although there is trace of compensation for the non-native type of change. We turn now to the L2-part of this study, presenting the results of American listeners hearing French sentences and of the French listeners hearing American English sentences.

3.2.3. Americans hearing French

One voicing item is excluded. Four participants from the beginner group (N=13) and one from the advanced group (N=8) had to be discarded. Compensation indices for both beginners and advanced learners are displayed in the right part of Fig. 2 above. The beginners have a mean compensation index for place of 45%, and of 16% for the voicing contrast, a significant difference by subjects and marginal by items ($F_1[1,12]=8.8$, $p<.012$; $F_2[1,29]=3.0$, $p>.05$). This difference indicates that beginning learners of French, when hearing French, do compensate more often for a place assimilation than for the French-native process, voicing assimilation. For the advanced group in contrast, the index for place is 22%, and 59% for voicing. The difference is significant by subjects and items ($F_1[1,7]=15.2$, $p<.006$; $F_2[1,29]=15.8$, $p<.0001$). Hence they behave like the French native speakers, and similarly to the advanced learners tested in experiment 1.

Unlike in experiment 1, comparing directly the behavior of the participants in French and in American English is now possible. Analyses include only those subjects who participated successfully in both experiments, omitting (i.e. from the complete American group also) those 5 subjects who have been excluded here. A by-subject ANOVA on indices, restricted to each proficiency level, shows a significant interaction between contrast and test language only for the advanced group ($F_1[1,14]=30.1$, $p<.0001$), but not for the beginner group ($F_1[1,24]=0.5$, $p>.4$). This confirms that beginners display the same pattern for French as for American English, whereas advanced learners changed their behavior in French as compared to American English. As appears from both studies involving beginners so far, native language phonology seems to count more than the phonology of the sentences. By contrast, more advanced learners did show evidence of having adopted at least in part the phonology of the non-native language when they listen to sentences in that language.

We then compare both groups on L1, in order to see if L2-proficiency influences the processing of L1. We performed a by-subject ANOVA on indices in L1 (American English) declaring a new factor: "proficiency in L2", as well as contrast. While contrast did show a main effect ($F_1[1,19]=47.0$, $p<.0001$), there was no effect of proficiency ($F_1[1,19]=1.4$, $p>.2$) and no interaction with the other factor ($F_1[1,19]=0.01$, $p>.8$). This means that globally, proficiency in L2 does not affect the pattern of results obtained for L1, but is in fact responsible for the different behavior observed in L2 seen in Fig. 2. We now turn to the results for French listeners hearing American English.

3.2.4. French hearing American English

The same four items that had been excluded for the American group (3.2.1.) were also removed here. Fourteen participants did not reach inclusion, leaving for analysis a group of 16 subjects, 8 beginners and 8 advanced learners. Compensation indices for beginners and advanced learners are shown in the right part of Fig. 3 above. Beginners' compensation index is 24% for the place contrast, and 62% for the voicing contrast. This difference is significant by subjects, marginal by items ($F_1[1,7]=6.1$, $p<.05$; $F_2[1,26]=3.7$, $p=0.06$). As we previously observed in other beginner groups, compensation in L2 is higher for the L1-process than for the L2 one. The advanced learner group obtained a compensation index of 36% for the place contrast and of 29% for voicing, but this difference failed to reach significance ($F_1[1,7]=0.6$, $p>.4$; $F_2[1,26]=0.5$, $p>.4$). Nevertheless, this pattern is different from that of the beginners: an ANOVA on indices for both groups declaring the factors proficiency (beginner vs. advanced) and contrast (place vs. voicing) found no main effect of type (marginal only by subjects, $p>.05$, $p>.3$), and no main effect of proficiency ($p>.2$, $p>.8$), but the significant interaction between both factors ($F_1[1,14]=6.5$, $p<.03$; $F_2[1,26]=5.4$, $p<.03$) reveals that proficiency does play a role in a differential compensation as a function of contrast. We restrict the analysis to one contrast type (first

place then voicing) to assess whether proficiency modifies the compensation for both types of contrast equally. No difference between beginners and advanced learners is visible on the place contrast (24% vs. 36% respectively, $F_1[1,14]=0.8$, $p>.3$; $F_2[1,12]=2.2$, $p>.1$). However, compensation on the voicing contrast seems to be more affected by proficiency, as the difference on indices is significant between beginners (62%) and advanced learners (29%) ($F_1[1,14]=8.1$, $p<.013$; $F_2[1,14]=3.3$, $p>.05$), although marginally by item.

Next, we compare the performance of the same participants on L1 and L2. We include only those who were successfully tested on both languages, removing 14 from the French results who had been excluded on American English and removing 1 beginning learner who had been excluded on French, leaving a total of 15. By-subject ANOVAs show for the advanced group only a significant interaction between contrast and test language ($F_1[1,14]=8,0$, $p=.013$), not for the beginner group ($F_1[1,12]=1.6$, $p>.2$). This confirms that beginners display the same pattern for American English as for French, whereas advanced learners changed their behavior in American English as compared to French, even though they did not completely switch their compensation behavior. However, the significant interaction between both factors confirms that for these French advanced learners also, test-language influences the compensation pattern for place vs. voicing.

We finally compare beginners and advanced on L1, in order to see if L2-proficiency influences the processing of L1. We performed a by-subject ANOVA on indices in L1 (French) declaring the factor "proficiency in L2", as well as contrast. While contrast did show a main effect ($F_1[1,13]=30.0$, $p<.0001$), there was no effect of proficiency ($F_1[1,13]=1.4$, $p>.2$) and no interaction with the other factor ($F_1[1,13]=0.5$, $p>.4$). This means that globally, proficiency in L2 does not affect the pattern of results obtained for L1.

4. General Discussion

Previous work has shown that perceptual compensation for phonological assimilation is driven by language-specific phonological knowledge (Darcy et al, under review; Koster, 1987; Otake et al., 1996; Beddor & Krakow, 1999). In this study, we explored compensation for assimilation in late learners of a second language, as a function of amount of exposure to the second language. The method and stimuli we used allowed us to compare the processing of an assimilation pattern that exists only in the native language, with one that exists only in the non-native language. We thus obtained a measure of the extent to which subjects listening to their second language were using the pattern of compensation appropriate to that L2, or were using instead the compensation appropriate to their native language.

Results of Experiment 1 with British English late learners of French show that although beginning learners inappropriately apply the compensation of their L1 to their L2, advanced learners have almost completely acquired the compensatory processes of the L2. Experiment 2 used a fully crossed design with American English learners of French and French learners of American English. Both groups applied the compensation pattern of their L1 onto L2 if they were beginning learners (mean exposure less than 1 year). However, more advanced learners started to significantly shift their pattern of compensation towards that of the L2. That is, American learners of French compensated more for voicing than for place assimilations (just like the native French did). Vice versa, French learners of English dropped their compensation for voicing, although their compensation for place did not reach the pattern of the native American English speakers. Interestingly, all late learners continued to apply the appropriate pattern of assimilation when they were hearing their own native language.

One might argue that the differences we found between beginners and advanced learners are due to a difference in vocabulary size. Indeed, knowing that "rop" is not a real French word might induce advanced participants to interpret it more readily as the target "robe" than beginners, who might be uncertain whether "rop" is a different word. However, such a lexical bias would apply equally for place and voicing items, and irrespective of the appropriate versus inappropriate context. In other words, given our design, lexical knowledge can only create a pattern of error rate, not a change in the compensation index. In fact, Darcy (2003) found similar patterns of assimilation for real words like "robe" and non-words like "nobe", showing that lexical knowledge is not involved here.

In brief, the combined results of these experiments suggest that within a few years of exposure to a second language, learners can build a separate system for the processes of L2, without modifying the L1 system, and are able to switch from one to the other depending on the language they are hearing. This conclusion contrasts with several claims that have been made regarding the plasticity of language-specific perceptual processes.

First, our findings contrast with claims of strong limits in the plasticity of perceptual processing. Non-native contrasts are difficult to perceive, and remain processed in a non-native way even after extensive training (e.g. Hallé et al., 1999; Iverson et al., 2003). Pallier, Bosch & Sebastián-Gallés (1997), Dupoux et al. (submitted) go further and claim that under naturalistic exposure conditions, a second language basically does not alter the way in which speech sounds are parsed onto native phonetic categories. Flege (1995; Flege, Bohn & Jang, 1997) has a less extreme position, and proposes that there is some plasticity allowing bilinguals to shift native category-boundaries as a function of L2 usage (see also Norris, McQueen & Cutler, 2003). Yet, these changes are small compared to the ones we reported. Notice however the difference: in our study, we tested French and English with stimuli

using phonemic categories that exist in both languages. We did not test how phonetic differences between the realization of voiced and voiceless stops may be perceived across the two languages; rather, we looked at what kind of compensatory mechanisms may exist when a lexical form is changed by a contextual process. This suggests that parsing continuous speech into phonetic categories on the one hand, and matching phonetic surface forms onto underlying lexical forms on the other, are performed by different computational systems, with different potentials for modification by linguistic experience[2].

Second, our findings contrast with claims that there can only be one phonological system, and that any potential acquisition is bound to modify the processing of the L1 as well as of the L2. For instance, Flege (1995) proposes that speech categories are flexible but common to L1 and L2. Peperkamp, Dupoux and Sebastian-Gallés (1999) found that more than 30% of the native French-Spanish bilinguals from mixed families display the 'stress deafness' effect typical of French monolinguals. This suggests that even massive exposure to two languages from birth is not sufficient to insure that the two languages will be processed in a native-like fashion. Here, instead, we found that participants can acquire a secondary system of phonological compensation rather smoothly without loosing the system applying in their native language. How do we account for this discrepancy? We could appeal to the same line of reasoning as the one proposed above, namely that the studies reporting evidence in favor of a single phonological system all dealt with perception or production of phonetic categories, whereas we consider a rule system that presumably deals with sounds that have already been categorized and focuses on the mapping of phonetic forms onto the lexicon. It is possible that such a mechanism is quite distinct from the one that performs categorization, in that it is much more plastic, and allows for the acquisition and use of multiple phonological systems.

In fact, there are independent reasons to propose that the mapping between surface and underlying forms requires a system that is both plastic and allows for multiple instantiations. Indeed, most often, we hear words that are phonetically distorted compared to the standard form (e.g. through foreign and regional accents or telephonic equipments). Yet, listeners learn to adapt quickly to synthetic speech (Greenspan, Nusbaum & Pisoni, 1988) and to artificially time-compressed speech (Dupoux & Green, 1997), whereby this adjustment transfers to novel words, sentences or speakers. In order to account for these results, it is necessary to postulate a system specifying how abstract word forms can be matched to a speech signal that departs in significant ways from the normal pattern. As a speculation, we propose that the system allowing phonological compensation for assimilation is the same that allows one to learn and compensate for specific patterns of pronunciation distortions, as they occur in

[2] Error rate analyses in our data show consistently more errors on L2 but no effect of proficiency, suggesting that categorization is still more difficult in L2, and less subject to improvement with more exposure.

synthetic speech or dialectal variants of one's native language. Further research is needed to explore this hypothesis in more detail.

Before closing, we would like to comment briefly on two minor aspects of our results. First, we observed that all groups did show a limited compensation for the non-native process respectively, even though it is less than for the native process. This suggests the existence of a language-independent mechanism of contextual contrast sensitivity, like the Feature Parsing mechanism proposed by Gow (2003), which would apply *in addition* to the language-specific compensation mechanism (see Darcy et al., under review, for further discussion).

Second, we uncovered an unexpected asymmetry between French learners of English and American learners of French. The progressive differentiation of the L2 system for French was not as complete as for American advanced learners. French learners seem to have more difficulties in acquiring the non-native process; indeed, they mostly inhibited compensation for voicing assimilation when hearing English, but did not increase their compensation for place to the extent that native English listeners did. One possibility is that voicing and place assimilation differ in the first place. There is phonetic evidence that French voicing assimilation is more categorical than English place assimilation (Féry, 2003; Nolan, 1992). This appears in the respective L1-data of the French and Americans. Voicing seems to be compensated for more completely by the French than Place by the Americans. Our stimuli with deliberate substitutions are then perhaps closer to real French voicing assimilation than to English place assimilation. This could explain the lower compensation for place assimilation among both English and French listeners. Similarly, because of such a difference between the French and English assimilation processes in natural speech, French learners of English have less clear evidence for the existence of place assimilation. Another possibility could be that, similarly to phonetic categories, it may be easier to adjust the weighting of assimilations that already occur even to a small extent in the L1 (as partial phonetic devoicing), than to establish a phonological category or process that does not exist at all in the L1. In this case, it would be harder for French to acquire place, because they have no familiarity at all with alternations modifying the place feature, whereas Americans would already have some sort of familiarity with variation along the voicing feature, due to the presence of this partial phonetic devoicing process. [3]

To recapitulate, as far as compensation for assimilation is concerned, we conclude that it is possible for the same person to have two distinct coexisting phonological systems and to switch from one to the other as a function of the heard language. This suggests that compensation requires mechanisms

[3] We are grateful to an anonymous reviewer for pointing this out.

different from those involved in phonetic categorization. More research is needed to determine the nature of these mechanisms.

References

Beddor, P. S. and Krakow, R. A. (1999). Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *Journal of the Acoustical Society of America 106*(5), 2868-2887.

Best, C.T (1995). A direct realist perspective on cross-language speech perception. In W. Strange and J. Jenkins (eds.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Timonium, MD: York Press.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. In D. Birdsong (ed.), *Second language acquisition and the Critical period hypothesis* (pp. 133-159). Mahway, NJ: Erlbaum.

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

Darcy, I. (2002). Online processing of phonological variation in speech comprehension: The case of assimilation. In *Temporal integration in the perception of speech*, S. Hawkins and N. Nguyen (eds.), 32. ISCA.

Darcy, I. (2003). Assimilation phonologique et reconnaissance des mots. Ph.D. Dissertation, Ecole des Hautes Etudes en Sciences Sociales Paris and Johannes Gutenberg Universität Mainz.

Darcy, I., Ramus, F., Christophe, A., Kinzler, K., and Dupoux, E. (under review). Phonological knowledge in compensation for native and non-native assimilation.

Dupoux, E., and Green, K. (1997). Perceptual adjustment to highly compressed speech: Effects of talker and rate changes. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 914-927.

Dupoux , E., Sebastián-Gallés, N., Navarrete, E., and Peperkamp, S. (submitted). Persistent stress deafness: the case of French learners of Spanish.

Ellis, L., and 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 *Nouveaux départs en phonologie,* T. Meisenburg, and M. Selig (eds.). Tübingen: Narr.

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In *Speech perception and language experience: issues in cross-language research,* W. Strange (ed.), 233-277. Baltimore, MD: York Press.

Flege, J. E., Bohn, O. S., and Jang, S. (1997). Effects of experience on non-native speaker's production and perception of English vowels. *Journal of Phonetics 25*, 437-470.

Flege, J. E., Frieda, E. M., and Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics 25*, 169-186.

Flege, J. E., Yeni-Komshian, G. H., and Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language 41*, 78-104.

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

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

Greenspan, S. L., Nusbaum, H. C., and Pisoni D. B. (1988). Perceptual learning of synthetic speech produced by rule. *Journal of Experimental Psychology. Learning Memory and Cognition 14*, 421-33.

Hallé, P. A., Best, C. T., and Levitt, A. (1999). Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics 27*, 281-306.

Hyman, L. M. (1975). *Phonology : Theory and analysis*. New York : Holt, Rinehart and Winston.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., and Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition 87*, B47-B57.

Johnson, J. S., and Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology 21*, 60-99.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language 60*, 286-319.

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

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics 50*, 93-107.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition 12*, 251-285.

McAllister, R., Flege, J. E., and Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics 30*, 229-258.

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

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

Norris, D., McQueen, J. M., and Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology 47*, 204-238.

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

Pallier, C., Bosch, L., and Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition 64*, B9-B17.

Peperkamp, S., Dupoux, E., and Sebastián-Gallés, N. (1999). Perception of stress by French, Spanish, and bilingual subjects. *Proceedings of EuroSpeech '99, Vol. 6*, 2683-2686.

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

Strange, W., Yamada, R. A., Kubo, R., Trent, S.A., Nishi, K., and Jenkins, J.J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics 26*, 311-344.

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

Weber-Fox, C., and Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience 8*, 231-256.

Author Notes

Isabelle Darcy

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

Universität Potsdam, Institut für Linguistik, Potsdam, Germany

Sharon Peperkamp

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

Université Paris 8 Saint Denis, France

Emmanuel Dupoux

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

Appendix

Detection rates (%) in each condition as a function of the type of contrast (place vs. voicing) with indices for each group.

	Place				Voicing			
Experiment 1	viable	Unviable	no-change	index	viable	unviable	no-change	index
English subjects hearing English sentences, n = 27	60	46	89	35	25	20	92	6
English subjects (low) hearing French sentences, n = 23	58	33	82	55	53	38	87	30
English subjects (high) hearing French sentences, n = 14	32	12	84	28	70	36	88	65
Experiment 2								
American subjects hearing American sentences, n = 26	46	11	94	43	33	18	91	20
American subjects (low) hearing French sentences, n = 13	49	24	84	45	32	22	84	16
American subjects (high) hearing French sentences, n = 8	25	6	87	22	61	26	86	59
French subjects hearing French sentences (n=23)	40	17	91	32	69	20	92	70
French subjects (low) hearing American sentences (n=8)	37	19	87	24	64	31	87	62
French subjects (high) hearing American sentences (n=8)	45	23	86	36	54	37	92	29