special is *especial* but *stuto* is not *astuto*: perception of prothetic /e/ in speech and print by speakers of Spanish*

Pierre A. Hallé, Juan Segui, Alberto Dominguez, and Fernando Cuetos

Abstract

In this presentation, we report several experiments on the "illusory" perception by native speakers of Spanish of a prothetic /e/ vowel in spoken nonwords or pseudowords beginning with /s/ followed by a consonant (e.g., **spid* or **special*, where '*' denotes phonological ill-formedness) as well as in 's'+C pseudowords presented in print (e.g., "special"). We begin with spoken nonwords such as **spid* derived from either *espid* or *aspid*, establishing the basic phenomenon of prothetic /e/ perception: Spanish listeners perceive **spid* as *espid*, that is, they perceptually repair **spid* into *espid*. French listeners do not repair *spid*, which is a well-formed sequence in French. We then gauge the consequences of this perceptual, prelexical repair for lexical access: Spanish listeners repair **stuto* (derived from *astuto*) into **estuto* more often than into *astuto*. That is, the phonological-level, prelexical repair prevails on a possible lexical-level repair. Finally, we use a visual masked priming paradigm to examine the processing of pseudowords such as "stuto" or "special" presented in subliminal print. We find that subliminal, printed "sC-" sequences are repaired into "esC-", just like their spoken counterparts. We therefore propose that the automatic, non-conscious processing of print does not produce lexical repair; it achieves not only phonological coding but also phonological repairs analogous to those observed for speech.

* [*special* es *especial* pero *stuto* no es *astuto*: la percepción de un /e/ protetico en palabras habladas o escritas]

This work has benefited from the ANR grant "PHON-REP" to the authors.

1. Introduction

About ten years ago, we serendipitously found that French listeners hear /kla/ when told /tla/, which is phonotactically illegal in French (Hallé, Segui, Frauenfelder, & Meunier, 1998). This case of misperception proved to be very robust in subsequent research. We assumed it was yet another case of foreign sound misperception, similar to those described in the early thirties by Polivanov (1931) for either single sounds or sound combinations. For example, Japanese listeners cannot distinguish /r/ from /l/. Polivanov labeled these cases of misperception as "phonological deafness." More recently, Catherine Best reformulated these cases of misperception as "perceptual assimilations" of nonnative speech (Best, 1995): foreign, nonnative sounds are interpreted in terms of native phonemic categories. We called the /tl/ → /kl/ perceptual substitution "phonotactic perceptual assimilation" because it concerns combinations of sounds rather than single sounds. We also call it "perceptual repair" because it "fixes" an illegal combination of sounds into a combination that is permissible in the listener's language.

More recently, we have explored vowel insertion perceptual repairs (i.e., perceptual vowel epenthesis). Such repairs have been well documented for instance in Japanese: Japanese listeners perceive */ebzo/ as /ebuzo/ (Dupoux et al., 1999). Spanish provides a similar though seldom studied case of vowel insertion: /e/ prothesis in word-initial /s/+consonant sequences. Native speakers of Spanish often pronounce foreign words such as French "statue" ('statue') /staty/ as /estatu/; foreign words such as English "snob" produce loans such as *esnob*. But do Spanish listeners actually *hear* an /e/ before /s/ in these forms? Because this perception issue had not yielded published report yet (with the exception of the MA dissertation of Theodore, 2003), we decided to explore it further and found that, indeed, Spanish-speaking listeners do hear /e/ in utterance-initial /s/+consonant clusters (henceforth,

#*sC* clusters), be they onsets of nonwords, such as **spid*, or of pseudowords derived from words, such as **stuto* derived from *astuto* (‘astute’).

This chapter does not report the experiments we conducted in chronological order but, rather, along a dimension of presumed representational complexity. We begin with a phonological, prelexical level of representation, using spoken nonwords such as **spid* in a vowel detection task, and then evaluate the possible influence of lexical representations, using spoken pseudowords such as **stuto* (from *astuto*) compared to **special* (from *especial*). We find that Spanish listeners massively “repair” #*sC* into #*esC* (e.g., **spid* into *espid*; **stuto* into *estuto*, **special* into *especial*), whether the repair produces a word (e.g., *especial*) or not (e.g., *estuto*). That is, lexical influences play quite a minor role: The #*sC* → #*esC* repair is dominantly phonological in nature, and prelexical in locus. Finally, we report visual masked priming experiments showing that this repair applies to printed sequences such as “stuto” or “special.” As has been found in many masked priming studies before, subliminal printed primes quickly “activate” a phonological code, which is presumably parallel to prime orthography. What is new in our studies is that phonological repairs (here, the #*sC* → #*esC* repair) apply to the phonological code activated from subliminal print.

The chapter is organized as follows: Section 2 presents the ban of #*sC* in Spanish from a diachronic and synchronic point of view. Sections 3 to 5 present our experimental data on the perception by Spanish listeners of #*sC* items in a low-level phonetic decision task (section 3), in an auditory lexical decision task (section 4), and in a visual masked priming task with lexical decision on visual target (section 5). The concluding comments bear on the nature and locus of the repair processes evidenced in this experimental work.

2. The ban of word-initial /s/+consonant in Spanish

Diachronically, a prothetic vowel before *#sC* appeared in Latin as an articulatory device facilitating the pronunciation of *#sC* immediately after a consonant. *C#sC* contacts themselves arose from the loss of word-final vowels in Low Latin. Zink (1986) proposed that the reinserted prothetic vowel provides an articulatory "appui" (a 'base') for /s/, and that it is "… articulée au plus près de s …" [articulatorily closest to /s/]. In Latin, the prothetic vowel before *#sC* thus could only be short /i/, which evolved into long /i/ or into /e/ in Romance languages. In 11th century old French, a prothetic vowel before *#sC* occurred only after a consonant (e.g., "il out espusethe" [he had married]), not after a vowel (e.g., "la spusa" [the spouse]). In 12th century old French, however, the prothetic vowel generalized, as well as in Iberian Romance languages. Later on, due to borrowings from old Latin, Germanic, and other languages, the systematicity of the prothetic vowel was lost in French, but not in languages such as Spanish. Two extreme situations can now be found in Romance languages with respect to *#sC* cognates from Latin: *#sC* is systematically maintained in Italian (e.g., Latin *schŏla* > Spanish *escuela* vs. Italian *scuola*), whereas it is banned in Spanish (or other Iberian Romance languages), in which it *always* changed to *#esC*. Spanish loanwords borrowed from languages that have maintained *#sC* undergo the same change: For example, English *snob* is adapted as *esnob*. Note that the other languages that ban *#sC*, also "repair" *#sC* with a prothetic vowel, which may be different from /e/: Catalan uses *schwa* (Bonnet & Lloret, 1998) and Brazilian Portuguese uses /i/ (or sometimes /u/) (Câmara Jr., 1969).[1] Cross-linguistically, prothesis in sibilant+consonant clusters is largely preferred to other repair strategies such as cluster simplification or within-cluster epenthesis (see Fleischhacker, 2001).

[1] The prothetic, or more generally, epenthetic vowel inserted in loanwords varies across languages. On a phonological account, it is motivated by markedness considerations: The borrowing language tends to choose its less marked vowel as the epenthetic vowel (e.g., /u/ in Japanese, the vowel most prone to deletion in phonological alternations). On a phonetic account, epenthetic vowels are the shortest and the most centralized vowels in a given language's phonetic space.

3. Detection of /e/ or /a/ in #*sC*-#*esC* and *sC*-#*asC* continua

In the early thirties, Polivanov (1931) proposed, on the basis of informal observation, that Japanese listeners repeat *drama* as *dorama* (or *dzurama*) *because* they do hear *dorama* in *drama*. Experimental work (Dupoux et al., 1999) confirmed Polivanov's intuition. Japanese listeners indeed heard an epenthetic vowel /u/ in all the stimuli of continua such as [ebuzo]-[ebzo] (constructed in gradually deleting [u]), even at the [ebzo] endpoint. In contrast, French listeners perceived the continua in a categorical way, from near-ceiling detection of /u/ at the [ebuzo] endpoint to no detection of /u/ at the [ebzo] endpoint. Japanese listeners thus *perceptually* repaired the illegal form [ebzo], complying with Japanese phonotactic constraints. For Spanish, Theodore (2003) similarly reported that Spanish listeners hear /e/ in all the stimuli of the [estib]-[stib] continuum, even at the [stib] endpoint.

The study we report in this section used the same design as in the two studies just mentioned, with some improvements (see Cuetos et al., 2011, for more detail). One concern in Dupoux et al.'s (1999) study, which applies as well to that of Theodore, is that digitally deleting a vowel in a speech sequence may leave traces of it in the surrounding segments. In the present study, the design was such that a prothetic vowel percept could not arise from traces of the vowel in question remaining after editing speech signal. This was achieved in using both #*esCid* and #*asCid* forms as base forms from which #*VsCid*-#*sCid* (V=[e] or [a]) continua were constructed. Testing the detection of /e/ in stimuli derived from #*asCid* vs. #*esCid* forms could indeed provide an estimation of how much possible traces of the original vowel would affect performance. In addition, listeners were also tested on their detection of /a/ instead of /e/ in the same stimuli. This allowed for testing the possibility that listeners be insensitive to vowel timber in the #*sC* endpoint region of the continua, and might give, at least in this region, false positive detection responses to *any* target vowel. We used a classic cross-language design, comparing Spanish to French listeners, who served as a reference for /e/ detection rates presumably unbiased by perceptual repair.

Forty-nine Spanish listeners from Oviedo (Spain) and 50 French listeners from Paris participated in theses studies. Each *#VsCid-#sCid* continuum (V = /a/ or /e/ and C = /p/, /f/, or /m/) consisted of six stimuli constructed from *#esCid* or *#asCid* items spoken by a native speaker of Standard Spanish. The first stimulus in each continuum was constructed in digitally deleting the entire item-initial /e/ or /a/ vowel. The other stimuli were constructed by progressively preserving a longer portion of the initial vowel, with a step of two voicing periods (15 ms in average), until the sixth stimulus, which was the unaltered *#VsCid* “base” item. Half of the participants had to detect /e/ in all these stimuli. The other half had to detect /a/. Spanish listeners detected /e/ in the first two stimuli about 56% and 30% of the time, regardless of whether the base item was beginning with /a/ or /e/, and regardless of whether C was /p/, /f/, or /m/. They detected /a/ only in the stimuli based on *#asCid*, and only from the second stimulus of the continuum onward, reaching a near-ceiling level of detection at the third stimulus. French listeners’ performance on /a/ detection was nearly identical to that of Spanish listeners. French listeners detected /e/ only in the stimuli based on *#esCid* just like they detected /a/ only in the stimuli based on *#asCid*. In other words, contrary to Spanish listeners, French listeners were not biased to detect /e/ more than /a/ in the first two or three stimuli of the continua. Figure 1 shows the Spanish listeners detection performance.

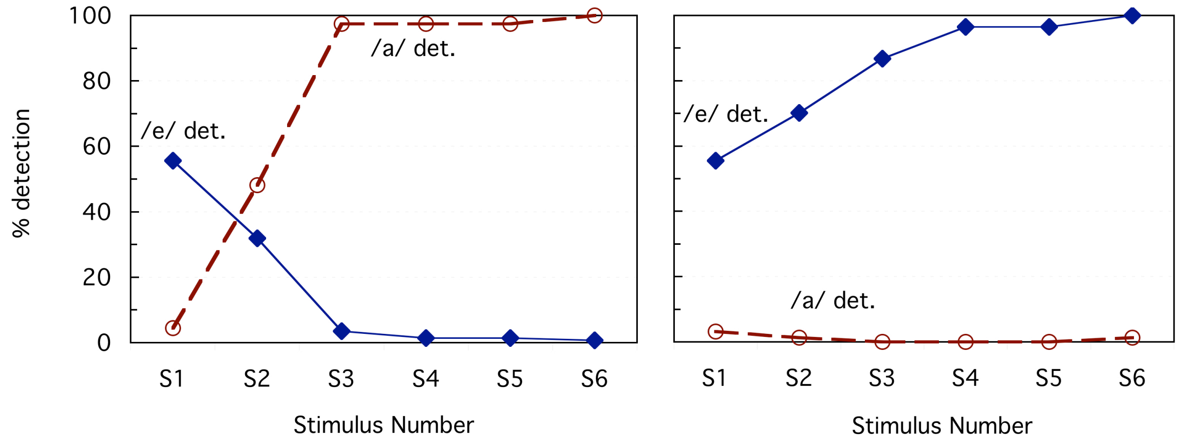

Figure 1. Spanish vowel detection data as a function of stimulus number (S1-6) for the stimuli based on *#asCid* (left panel) or *#esCid* (right panel) items.

Lexical effects were unlikely in this task. The results thus should be mainly interpreted as documenting a purely or at least largely phonological perceptual repair in Spanish. The design we used moreover allowed for a few qualifications on this repair. First, it is not induced by traces of the deleted vowel: Even for stimuli with about two voicing periods left of the word-initial [a], the rate of /e/ detection is substantial (>30%). Second, listeners' performance for /a/ detection showed excellent sensitivity to vowel timber, however short the vowel was. We therefore assume that the results were free from a bias for false positive response. In the next section, we turn back to the issue of lexical effects, using a design that allows for the estimation of the relative contributions of phonological and lexical repairs.

4. Auditory lexical decision on **stuto* versus **special*

A previous study (Hallé & Segui, 2003) conducted with Argentinian listeners suggested mixed effects of lexical feedback and prelexical repair in a free transcription task on auditory *#sC* stimuli. These *#sC* stimuli were pseudowords derived from either an /e/ word such as *especial* (**special*) or an /a/ word such as *astuto* (**stuto*). Listeners faithfully transcribed the *#sC* stimuli as beginning with /s/ about 66% of the time for both **special* and **stuto*. For **stuto*, 13% responses were lexical repairs (i.e., “astuto” for **stuto*) and 21% were phonological repairs, that is, responses beginning with ‘e’ (i.e., “estuto” for **stuto*). For **special*, 34% responses were ‘e’ responses. Although the lexical vs. phonological repair status of the latter responses logically cannot be decided, the fact that 21% phonological and 13% lexical for **stuto* exactly add up to 34% is consistent with a similar distribution for **special* and **stuto*. This preliminary transcription study therefore suggested that about 62% (21/34) of the repairs are phonological in nature (i.e., prelexical), whereas 38% are lexical.

There might be biases of many kinds in the free transcription study just described: In particular, responses might be influenced by the written forms of English words known by the subjects. We therefore decided to go for a more controlled experimental setting whereby phonological and lexical contributions to the *#sC* → *#esC* repair could be estimated. We used similar materials to those in the transcription study and submitted them to auditory lexical decision by Spanish listeners. The critical comparison was, again, between pseudowords such as **stuto* and **special*, with the idea that “word” responses to **stuto* correspond to lexical repair or to false alarm, and “word” responses to **special* correspond to either lexical or phonological repair or to false alarm. Assuming that similar rates of lexical repair and false alarm arise for both types of **stuto* and **special* pseudowords, the “word” response rate difference between **special* and **stuto* would estimate the rate of phonological, prelexical repairs. We complemented these critical materials with the unaltered word forms corresponding to the *#sC* pseudowords (e.g., *astuto* and *especial*), as a reference for optimal “word” response rate. Were the “yes” response rate to **special* smaller than that for *especial*, the difference would be an estimation of correct rejection rate for **special*.

Figure 2A shows the rate of "word" responses obtained with 41 Spanish listeners (from Oviedo, Spain). Figure 2B shows the corresponding response times (RTs), which were measured, for all the test stimuli, from the release of the stop following /s/. Unaltered *especial* and *astuto* words yielded nearly identical "word" response rates and RTs; significantly lower rates and longer RTs obtained for **special* pseudowords, and still lower rates and longer RTs obtained for **stuto* pseudowords.

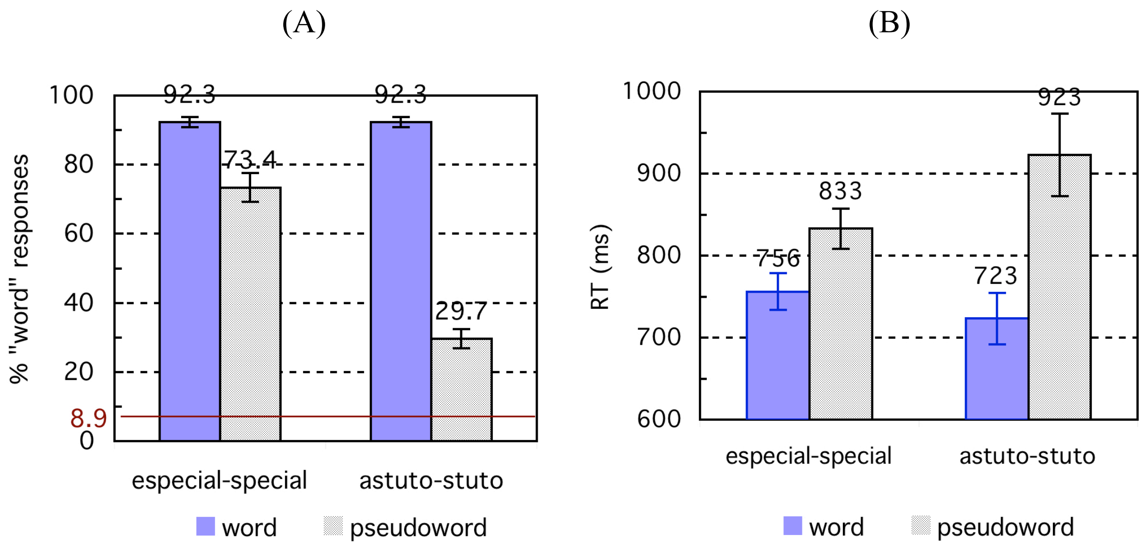

Figure 2. Auditory lexical decision data for *astuto* vs. *especial #VsC* words and the derived *#sC* pseudowords: (A) rate of "word" responses, and (B) response times.

We may now estimate quantitatively the lexical and phonological repair effects from the "word" response rate data in the following way: The average rate of "word" responses to nonword fillers (8.9%) provides an estimate of false alarm rate in the absence of lexical bias; we may thus consider that the 29.7% "word" responses to **stuto* pseudowords decompose into 9.6% false alarms and 29.7-8.9≈20.8% "lexical repair" responses; pursuing this line of reasoning, the 73.4% "word" responses to **special* pseudowords decompose into 73.4-29.7≈43.7% "phonological repair" responses, and the same 8.9% false alarm and 20.8% "lexical repair" responses. To sum up, repairs would be approximately distributed into 69%

phonological and 31% lexical repairs. This estimation is consistent with that from the earlier transcription study (Hallé & Segui, 2003). Importantly, phonological repairs are dominant: They are twice as frequent as lexical repairs.

The RT data reveal another important aspect of the observed repairs: They have a cost in terms of processing time. Indeed, RTs for "word" responses were longer for **special* pseudowords than *especial* words (833 > 756 ms), and were longer still for **stuto* pseudowords compared to *astuto* words (923 >> 723 ms). (Note that these raw figures could be refined by taking into consideration the distributions of false alarm, lexical repair, and phonological repair "word" responses to **stuto* and **special*.) These RT data lead to the speculation that lexical repairs, which require substantial additional processing time (about 200 ms), are underlain by strategic, conscious processes. In contrast, phonological repairs, which only require a moderate amount of additional processing time (less than 100 ms), may be accomplished through more automatic, on-line processes. In the visual masked priming study presented to close this chapter, we address the question of whether at least part of the repair process is automatic, mechanical as it were, and operates non-consciously.

5. Phonological repair from subliminal print

Numerous studies showed that subliminal print activates phonological representations, that is, a phonological code. Phonological code from print seems to build up somewhat later than orthographic code (Ferrand & Grainger, 1993) but seems to do so automatically, out of conscious control. Now, within the visual masked priming paradigm, whereby printed primes are subliminal (non-consciously perceived), would "special" non-consciously activate a phonological code such as **special*, which would then be repaired into *especial* and prime "especial"? Would "stuto" likewise activate **stuto*, which would then be repaired into *astuto* and prime "astuto"? If the hypothesis that concluded the previous section is correct, only phonological repairs but not lexical repairs may operate non-consciously. On this hypothesis, "special" should prime "especial" but "stuto" should not prime "astuto." This is indeed what we found in a series of experiments reported in full detail in Hallé et al. (2008).

In this chapter, we only summarize this series of experiments. Printed targets such as "ESPECIAL" vs. "ASTUTO" were presented after primes such as either "especial" vs. "astuto" (repetition condition) or "special" vs. "stuto" (alteration condition). Primes were always presented for 44 ms. Prime was followed by the target immediately (44 ms SOA), after a 44 ms mask (88 ms SOA), or after a 44 ms mask and a 44 ms blank screen (132 ms SOA). The participants' task was a lexical decision on the target, which was maintained on the screen until participant's response. In this design, the repetition condition was the reference condition for which maximal priming, hence minimal RT, was assumed. Longer RTs in the alteration than the repetition condition would indicate a lesser priming, whereas similar RT would indicate similar priming, thus presumably indicating repair of the altered prime. The alteration effect (the RT difference between the alteration and repetition conditions for a given target) was thus the critical outcome in these experiments. Figure 3 shows that the alteration effect was negligible at 44 ms SOA for both the "stuto" and "special" primes and became substantial (~50 ms) at longer SOAs for the "stuto" primes only; for the "special" primes, the alteration effect was moderate at 88 ms SOA (~20 ms) and became negative at 132 ms SOA (~ -20 ms). We interpreted these data as reflecting equivalent orthographic priming at the shortest SOA, and the emergence of phonological priming at longer SOAs.

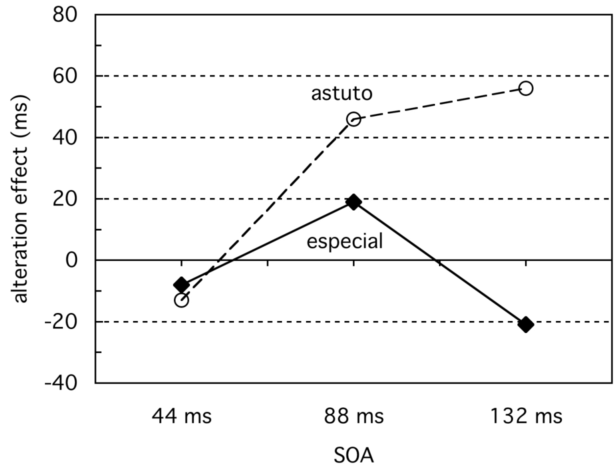

Figure 3. Alteration effects according to prime–target SOA: “astuto” vs. “especial” targets.

Because the size of the alteration effect found for the “stuto” primes at long SOAs was of the same order of magnitude as the largest phonological priming effects found with the visual masked priming paradigm, we may surmise that the phonological code activated by “stuto” was not compatible at all with “astuto.” In other words, this code was not lexically repaired into *astuto*. In contrast, the phonological code activated by “special” built up with increasing SOA and was apparently compatible with “especial.” This code was thus presumably repaired into *especial*. The “special” primes induced decreasing alteration effect with increasing SOA, until the alteration effect tended to reverse in favor of “special” rather than “especial” at 132 ms SOA. This reversal was interpreted as reflecting the additional processing time required for phonological repair, which presumably entailed an earlier occurring decay of “especial” facilitation by “especial” compared to “special.”

In this chapter, we would like to stress the difference in nature between lexical and phonological repair. The visual masked priming data show that repairs compatible with target apply to the “special” but not the “stuto” primes. This suggests that the repair processes operating on non-conscious primes can only be phonological repairs, not lexical repairs. Such a difference in nature between lexical and phonological repair is in line with the auditory lexical decision RT data (section 4), which suggested that lexical repairs (e.g., **stuto* > *astuto*) arise from conscious reevaluation of the overtly presented stimuli, whereas phonological repairs operate automatically. We propose that the *#sC* → *#esC* phonological repair discussed in sections 3 and 4 applies to both types of primes, "special" and "stuto." By this account, "special" and "stuto" would first generate phonological codes such as **special* and **stuto*, which would then be repaired into *especial* and **estuto*, explaining why "special" facilitates "especial" but "stuto" does not facilitate "astuto."

6. Concluding comments

Across three studies we found that native Spanish-speaking listeners "hear" a prothetic /e/ in ill-formed spoken utterances such as **spid* or **stuto*, as well as in ill-formed printed sequences such as "stuto." We interpret this phenomenon as a perceptual phonological repair, which turns illegal *#sC* into *#esC*. We claim that this phonological repair is prelexical in locus and automatic in nature. The first study (section 3) suggests it occurs at a low perceptual level. The second study (section 4) suggests that the *#sC* → *#esC* phonological repair is not inhibited during tentative lexical access of overtly presented illegal pseudowords: in an auditory lexical decision task, **stuto* seems to be perceived as the nonword *estuto* rather than the word *astuto*. Finally, the third study (section 5) suggests that the *#sC* → *#esC* repair takes place during the non-conscious processing of subliminal print. This repair operates on the phonological code activated from subliminal printed sequences. Moreover, no lexical repair takes place at this non-conscious level of processing. In other words, lexical repairs such as **stuto* → *astuto* are restricted to conscious reevaluation of overt stimuli.

Phonological repairs require additional processing time. This can be seen in the auditory lexical decision study, in which "word" response RTs to **special* are longer than those to unaltered *especial*. This can also be inferred from the visual masked priming study. First, phonological repair is clearly achieved only at SOAs between 88 and 132 ms, that is, probably later than is observed for simple phonological priming without repair as reported in the literature (see, for example, Ferrand & Grainger, 1993). Second, the reversal of the alteration effect at 132 ms SOA can be straightforwardly accounted for by a time-delayed time course of activation for "special" primes compared to "especial" primes.

The same *#sC* → *#esC* phonological repair applies to the phonological code activated from subliminal print and to overtly presented spoken items. This commonality between speech and print is clearly in line with the "unconventional motor theory" of Alvin Liberman, who proposed that print and speech activate a common phonological code, possibly grounded on articulatory phonetic specifications (Liberman, 1996). That the speech code activated from print may be processed in the same way as overt speech, with the same repairs taking place if needed, may also suggest that we internally "listen to" and monitor that speech code.

References

Best, C. (1995). A direct realist perspective on cross-language speech perception. In W. Strange, & J. Jenkins (eds.), *Cross-language speech perception* (pp.171–204). Timonium, MD: York Press.

Bonnet, E., & Lloret, M-R. (1998). *Fonologia catalana*. Barcelona: Editorial Ariel.

Câmara Jr., J.M. (1969). *Problemas de lingüistica descritiva*. Pertopolis: Vozes.

Cuetos, F., Hallé, P., Dominguez, A., & Segui, J. (2011). Perception of prothetic /e/ in #sC utterances: Gating data. In W.-S. Lee & E. Zee (eds.), Proceedings of *the 17th International Congress of Phonetic Sciences* (pp. 540-543). Hong Kong University Press.

Dupoux, E., Kakehi, Y, Hirose, C., Pallier, C., Mehler, J. 1999. Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1568-1578.

Ferrand, L. & Grainger, J. (1993). The time-course of phonological and orthographic code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society*, *31*, 119-122.

Fleischhacker, H. (2001). Cluster-dependent epenthesis asymmetries. In A. Albright & T. Cho (eds.), *UCLA Working Papers in Linguistics 7, Papers in Phonology 5* (pp. 71-116). Department of Linguistics, UCLA: USA.

Hallé, P., Dominguez, A., Cuetos, F., & Segui, J. (2008). Phonological mediation in visual masked priming: Evidence from phonotactic repair. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 177-192.

Hallé, P., & Segui, J. (2003, September). *Special is especial: Perception and production of the /sp, st, sk/ clusters by native speakers of Spanish*. Communication at the 13th ESCOP, Granada, Spain.

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.

Liberman, A. M. (1996). *Speech: A special code*. Cambridge, MA: MIT Press.

Polivanov, E. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague*, *4*, 79-96.

Theodore, R. M. (2003). *Perceptual prothesis in native Spanish speakers*. MA Thesis, Kent State University.

Zink, G. (1986). *Phonétique historique du français*. Paris: Presses Universitaires de France.