PERCEPTION OF PROTHETIC /e/ IN #*sC* UTTERANCES: GATING DATA

F. Cuetos[a], *P. Hallé*[b,c], *A. Domínguez*[d] & *J. Segui*[c]

[a]Universidad de Oviedo, Spain; [b]LPP-Paris 3, France; [c]LPNCog-Paris 5, France;
[d]Universidad de La Laguna, Spain
fcuetos@uniovi.es; pierre.halle@parisdescartes.fr; adomin@ull.es;
juan.segui@parisdescartes.fr

ABSTRACT

Perception of an epenthetic vowel within illegal, nonnative, clusters has been reported for a number of stimulus vs. listener language situations. In this study we examine the perception of a prothetic /e/ before word-initial #*sC* clusters by Spanish vs. French listeners, using gated stimuli derived from naturally produced /#as/+C and /#es/+C utterances. Spanish listeners but not French listeners heard a leading /e/ in the early gates where only a short portion or no trace at all of the original vowel was left. The extent of this effect was not modulated by how much ill-formed the /s/+C sequences were in terms of universal sonority contour preferences.

Keywords: prothesis, anaptyxis, perceptual repair, Spanish, sonority cycle

1. INTRODUCTION

In the early thirties, Polivanov [11] proposed that Japanese listeners repeat the word *drama* as *dorama* (or *dzurama*) *because* they hear *dorama* in *drama*. This intuition based at the time on informal observation has been experimentally tested since then, in particular in the work of Dupoux's group. Japanese listeners indeed hear an epenthetic vowel /u/ within clusters such as [bz] [5]. The sequence [ebzo], impermissible in Japanese, is *perceptually* repaired into /ebuzo/, complying with the Japanese phonotactic constraints. Such repair is not due to lexical feedback: for example, [mikdo] is repaired into the nonword **mikudo* rather than the word *mikado* [6]. Since then, several studies converged to show that listeners perceptually repair nonnative phoneme sequences that are illegal in their native language into permissible sequences. Note that perceptual repairs do not always consist in vowel epenthesis [10]: For those languages in which /#tl/ is illegal, native listeners seem to repair /tl/ into /kl/ [8]. As for epenthesis repairs, the preferred type of vowel insertion (prothesis vs. anaptyxis) tend to follow the type of illegal cluster (sibilant + stop vs. obstruent + sonorant, respectively, cf. [7]). But most perceptual studies on vowel epenthesis focus on within-cluster vowel insertion (anaptyxis). Little attention has been paid so far to vowel prothesis perceptual repairs. Yet, Standard Spanish (as well as virtually all Iberic languages) is an ideal candidate for such studies because #*sC* clusters are illegal in Spanish and #*sC* foreign words regularly adapted into *#esC* words (e.g., English *snob* > *esnob*). In a study similar to the first experiment in [5], Theodore [13] reported that Spanish listeners hear /e/ in all the stimuli of continua such as [stib]-[estib] even at the [s] endpoint of the continua. One concern in Dupoux et al.'s study, which applies as well to that of Theodore, is that digitally deleting a vowel in a sequence may leave traces of it in the surrounding segments. In the present study, the design is such that a prothetic vowel percept could not arise from traces left by speech signal editing or unintended production of the vowel in question. This is achieved in using both *#esC* and *#asC* forms, from which are derived two series of gates in which the portion retained decreases from full item down to no initial vowel left. Testing the detection of /e/ in stimuli derived from *#asC* vs. *#esC* forms would provide an estimation of how much traces of the original vowel possibly left in the speech signal affect performance. Finally, in a classic cross-language design, we compare Spanish to French listeners, who serve as a reference for /e/ detection rates unbiased by perceptual repair.

Berent and colleagues [1] showed that native speakers of English hear two syllables in utterances such as *bnif, bdif*, or *lbif, all illegal in English*, and perceive them as *bə.nif, bə.dif*, or *lə.bif*. This epenthetic /ə/ effect was increasingly large from *bnif* to *bdif* to *lbif*. The authors proposed this graded effect reflected graded ill-formedness in terms of syllabic sonority contour [3]. As a secondary aim of the present study, we put to test this account in using three degrees of

sonority contour ill-formedness, similar to those in [1]: #*smid*, #*sfid*, and #*spid*.

2. GATING EXPERIMENT

Natural utterances of #*VsCid* items (V=/a/ or /e/) were gated backward from an edited stimulus in which the entire V and a portion of /s/ had been deleted, to the intact item stimulus. For each gate, French and Spanish listeners decided whether they heard an initial /e/ in one condition, or an initial /a/ in another condition.

2.1. Method

2.1.1. Participants

Forty-nine native speakers of Spanish, students at Oviedo University (35 female, 14 male; mean age 24 years), and 50 native speakers of French, students at Paris 3 or 5 (30 female, 20 male; mean age 25 years) participated in the experiment. Three additional French subjects were tested but their data were not retained (failure to complete the test in two cases, hearing non-initial /e/s in another case). The Spanish participants had had little exposure to French but some to English as University students. None of the French or Spanish participants reported hearing deficit nor any kind of language impairment.

2.1.2. Stimuli and design

The gating stimuli were constructed from six original disyllabic items: *espid, esfid, esmid, aspid, asfid, and asmid*. Two versions of each was used: one with a strong-weak (SW) stress pattern (e.g., *ESpid*), one with a weak-strong (WS) pattern (e.g., *esPID*). These items all are nonsense words in Spanish but are phonologically permissible: for example, /id/ is a possible rime in Spanish; its /d/ is more or less spirantized depending on regional accent. Four repetitions of each *VsCid* utterance were recorded by a female native speaker of Spanish from Extremadura, where /d#/ is usually half-spirantized into [ð]. Twelve utterances were retained on the basis of prosodic homogeneity and pronunciation clarity. From each of them, six gates were constructed after careful inspection of the last few periods of the vowel merging into /s/ in the following way: for the first gate, the initial vowel V was entirely deleted as well a few ms of /s/ beginning (We deleted /s/'s beginning until 3 ms after the spectral derivative function peaked at the boundary between V and /s/.); the remaining part of the utterance was ramped in with a 40 ms linear ramp; for the second gate, the last two or three periods of V were left; no ramping was applied; V was deleted until the sample closest to zero amplitude on a rising or falling portion of the waveform; for SW items, F0 in V ranged from 240 to 260 Hz; voicing periods were thus about 4 ms long; for WS items, F0 was in the [140, 160] range; voicing periods thus were 6-7 ms long; we chose to progress backward in the waveform by steps of 3 periods for SW, and of 2 or 2.5 periods for WS stimuli, that is, by about 12 ms steps in SW and about 14 ms steps in WS items; from gate 3 to gate 5, increasing portions of V were left in this way; finally, gate 6 was simply the intact original.

One stimuli list was constructed for each stress pattern. One list contained 36 stimuli (6 originals x 6 gates) repeated 4 times each, hence a total of 144 stimuli. Each list was assigned to two groups of participants: one group had to detect /e/, the other to detect /a/. There were thus four groups of Spanish and four of French participants, with 12 or 13 participants per group.

2.1.3. Procedure

Spanish and French participants were individually tested in sound-treated booths in Oviedo and Paris, respectively. The experiment was run using the Praat "ExperimentMFC" module [2]. Stimuli were presented in random order, at a comfortable listening level through professional headphones. On each trial, participants heard the stimulus just once, then had to decide whether or not it began with the target vowel /e/ for half the participants, or with /a/ for the others; participants then had to rate how well they believed their response fit the stimulus on a 1-5 scale. No time limit was imposed and participants were not allowed to revise their judgments once entered.

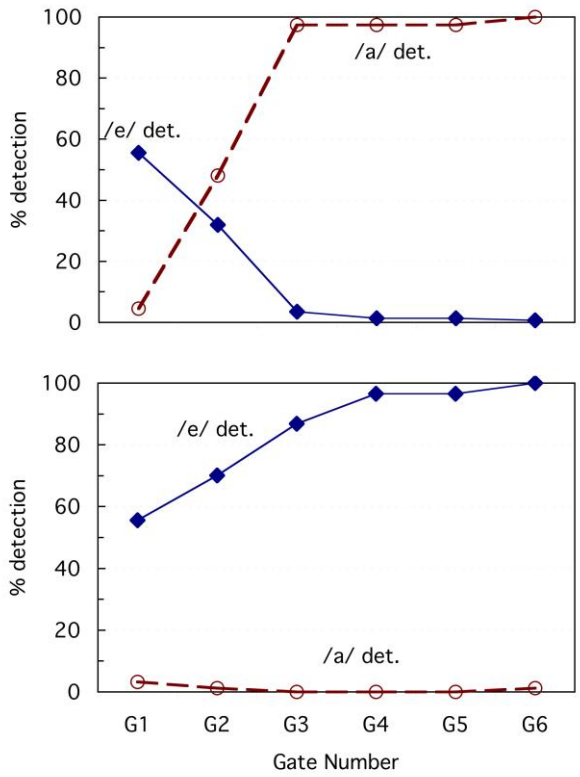
2.2. Results

Raw detection data for Spanish participants on the SW patterns are shown in Fig. 1, pooled across the Cs in *VsCid*. Closely similar results obtained with the WS pattern (see Table 1). We used the rating data to compute an "integrated" index of detection ((raw-rate x rating) / 5). Because raw detection rate and integrated index yielded essentially similar patterns of results we only report the raw rates.

As can be seen in Fig. 1, Spanish participants detected /e/ even at gate 1, about 56% of the time, whether V in the original stimuli was /e/ or /a/. On the following gates, /e/ detection rate increased

with the /e/ stimuli and decreased with the /a/ stimuli until near-ceiling detection with /e/ and near-floor detection with /a/ stimuli (cf. Table 1). At gate 2, where 12-15 ms of the original vowel were left, Spanish participants still detected /e/ in the /a/ stimuli at a substantial rate (32% for SW, 53% for WS patterns); at gate 3, /e/ detection rate dropped down to 3.5% for SW patterns but was still 23% for WS patterns, before approaching floor at gate 4 for both SW and WS. In contrast with /e/ detection, /a/ detection rate was near zero at gate 1 for /a/ or /e/ stimuli, then increasing up to near-ceiling for /a/ stimuli and remaining close to zero for /e/ stimuli.

Figure 1: Spanish vowel detection data as a function of gate number (G1-6); stimuli gated from SW *asCid* (top panel) or *esCid* (bottom panel).



The /a/ detection data are almost identical for French and Spanish participants. The French /e/ detection data in /e/ vs. /a/ stimuli also are near-identical to the Spanish or French /a/ detection data in /a/ vs. /e/ stimuli, respectively. The three latter conditions actually illustrate the expected patterns of vowel detection in backward gated stimuli beginning with either that same vowel or a different one, in the absence of perceptual repair bias. They all may serve as reference conditions to measure such a possible bias in the first condition, that is, detection of /e/ in /e/ vs. /a/ stimuli by Spanish listeners. In Table 1, part of the Spanish and French data (first four gates, and SW stimuli) shows target vowel detection rate, according to whether the stimulus vowel is the same or is different.

Table 1: Detection rate of /e/ vs. /a/ target according to whether the stimulus vowel is the same as the target vowel or not. The shaded portions correspond to the three "reference" conditions mentioned in the text.

V target to detect		/e/		/a/	
V' of the stimulus		*V=V'*	*V≠V'*	*V=V'*	*V≠V'*
Spanish subjects	G1	55.6	55.6	4.5	3.2
	G2	70.1	31.9	48.1	1.3
	G3	86.8	3.5	97.4	0.0
	G4	96.5	1.4	97.4	0.0
French subjects	G1	0.0	0.0	0.0	0.0
	G2	26.2	0.0	37.5	2.5
	G3	77.4	6.0	88.3	6.7
	G4	92.9	4.8	95.8	3.3

We ran an analysis of variance on the detection data with Subject as random factor and detection rate as the dependent variable. Between-subject factors were Language (Spanish vs. French), stress Pattern (SW vs. WS), and detection Target (/a, e/); within-subject factors were stimulus Vowel (same as Target vs. different), Consonant (/p, f, m/), and Gate number (1-4). (We restricted the analysis to the four first gates, given that ceiling or floor are reached at fourth gate in all conditions.) The main factors Language, Target, Vowel, and Gate were all significant at the $p<.00001$ level; stress Pattern and Consonant were not, Fs<1. The Language x Target interaction was highly significant overall, $F(1,91)=32.55$, $p<.00001$. The Spanish and French groups differed on /e/ detection, $F(1,45)=43.87$, $p<.00001$, not on /a/ detection, $F<1$. The other way round, Target was significant in the Spanish group, $F(1,45)= 50.33$, $p<.00001$, not in the French group, $F(1,46)=1.96$, $p=0.17$. The absence of a Language effect for target /a/ and of a Target effect for the French group reflect the strong similarity of the three shaded areas in Table 1, whereas the robust Language effect for target /e/ and Target effect in the French group reflect the different patterns of detection in the non-shaded area: detection of /e/ by Spanish listeners. In the shaded areas, detection of V in same-vowel stimuli increases with Gate number –from floor to ceiling rate– but remains at floor level in different-vowel stimuli. The pattern found in the non-shaded area is quite different. This difference is captured by the significant Gate x Language interaction for the /e/ target, $F(3,135)=$

46.62, $p<.00001$, but not the /a/ target, $F(3,138)= 1.04$, $p=0.38$. It can also be captured in more detail by focusing on the Spanish group, looking at the effect of Target for same- and different-vowel stimuli, according to Gate. In same-vowel stimuli, detection rate for /e/ is higher than for /a/ on the first two gates (ps<.0001), but becomes similar at gates 3 and 4 (Fs<1). In different-vowel stimuli, the same advantage for /e/ over /a/ obtains and remains significant at gates 3 and 4, ps<.01, though it has become numerically very weak.

Returning to the possible effect of sonority contour, that is, of the factor Consonant, no significant effect was found overall, F<1; it was not significant either in whatever subset of the data, for instance within the Spanish or the French data only, or within the strong-weak or the weak-strong data only, ps >.2.

3. DISCUSSION

Our data clearly establish perceptual vowel prothesis in Spanish for word-initial #sC clusters. Spanish listeners hear /e/ in *spid*, *sfid*, or *smid*. French listeners do not. The concern with possible traces in the acoustic signal of this perceived prothetic /e/ is directly addressed in our study: in stimuli where there are traces of *another* vowel (/a/ instead of /e/), Spanish listeners still hear /e/. Does this mean they might hear *any* vowel if requested to hear it? That they do not detect /a/ in *esCid* speaks to the contrary. Hence, our data provides a solid basis to ascertain that Spanish listeners hear /e/ in #sC at a phonetic, non-lexical level.

Sonority contour does not seem to modulate that trend, with contradictory, non-significant tendencies in the French data across stress pattern and between the French and Spanish groups. A possible reason for our failure to replicate Berent et al.'s findings is that #sC clusters have a special status cross-linguistically. Indeed, words beginning with /s/+C are quite frequent in English, French, or Italian, to only cite a few languages. Yet, monosyllabic words such as English (or French) *sport*, for example, are very common and all violate the sonority profile principle: their /sp/ onset has a falling sonority contour. Other possible analyses of such utterances consider them as *not* strictly monosyllabic: for instance, /s/ could be a syllabic consonant and *sport* be analyzed as *s.port*; or /s/ could be left unassigned to a syllable in the surface form tier: /s/ would then be an extrasyllabic consonant, a notion introduced by Clements and Keyser [4] (also see [12]), in line with that of core syllabicity [3].

Perception of a prothetic /e/ by Spanish listeners in #sC word-initial clusters has also been implicit in other previous studies using pseudowords derived from #esC vs. #VsC words (with V different from /e/: for example, *especial* vs. *astuto*) by deleting the initial vowel. These pseudowords are perceptually repaired by an epenthetic /e/ whether or not this repair produces a word [9]. Yet, this conclusion is somewhat indirect since [9] used a visual masked priming experimental paradigm. The data reported in the present study directly taps on auditory perception to demonstrate prothetic perceptual repair.

4. REFERENCES

[1] Berent, I., Steriade, D., Lennertz, T., Vaknin, V. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104, 591-630.

[2] Boersma, P., Weenink, D. 1993-2010. Praat: Doing phonetics by computer. *http://www.fon.hum.uva.nl/praat/*

[3] Clements, G.N. 1990. The role of the sonority cycle in core syllabification. In Beckman, M., Kingston, J. (eds.), *Papers in Laboratory Phonology 1*. Cambridge: Cambridge University Press, 283-333.

[4] Clements, G.N., Keyser, S. 1983. CV phonology: A generative theory of the syllable. *Linguistic Inquiry Monograph* 9. Cambridge, MA: MIT Press.

[5] Dupoux, E., Kakehi, Y., Hirose, C., Pallier, C., Mehler, J. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *J. Exp. Psych.: H. P. & P.* 25, 1568-1578.

[6] Dupoux, E., Pallier, C., Kakehi, Y., Mehler, J. 2001. New evidence for prelexical phonological processing in word recognition. *Lang. & Cog. Processes* 16, 491-505.

[7] Fleischhacker, H. 2001. Cluster-dependent vowel insertion asymmetries. *UCLA Working Papers in Linguistics: Papers in Phonology* 5, 71-116.

[8] Hallé, P., Best, C. 2007. Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *J. Acoust. Soc. Am.* 121, 2899-2914.

[9] Hallé, P., Domínguez, A., Cuetos, F., Segui, J. 2008. Phonological mediation in visual masked priming: evidence from phonotactic repair. *J. Exp. Psych.: H. P. & P.* 34, 177-192.

[10] Peperkamp, S. 2007. Do we have innate knowledge about phonological markedness? Comments on Berent, Steriade, Lennertz, and Vaknin. *Cognition* 104, 631-637.

[11] Polivanov, E. 1931. La perception des sons d'une langue étrangère. *TCLP* 4, 79-96.

[12] Rialland, A. 1994. The phonology and phonetics of extrasyllabicity in French. In Keating, P. (ed.), *Papers in Laboratory Phonology III*. Cambridge: Cambridge University Press, 136-159.

[13] Theodore, R., Schmidt, A. 2003. Perceptual prothesis in native Spanish speakers [Abstract]. *J. Acoust. Soc. Am.* 113, 2256.