Allophonic Variation Can Affect L2 Speech Perception: Evidence from a Tuscan Neutralization Process

Chiara Celata

Scuola Normale Superiore, Pisa, Italy
c.celata@sns.it

1. Introduction and purposes

The notion of perceived similarity between non-native sounds and corresponding native phonological categories is widely recognized as an important element for any account of L2 speech perception. In psycholinguistic models such as the SLM by Flege and colleagues and the PAM by Best and colleagues (Flege, 1987; Flege, 1995; Guion et al., 2000; Best, 1995; Best et al., 1988; Best et al., 2001; Best & Tyler, 2007; see also Munro & Bohn, 2007), the mapping of L2 sounds onto L1 speech categories is seen as the core of L2 perceptual processing. The ease or difficulty of a nonnative contrast depends on the perceived similarity to the corresponding sounds of the native phonology. For example, within the PAM (Best, 1995), different types of perceptual assimilation are individuated, based on the probability a non-native speech contrast has to be assimilated to a native phonological contrast. If an L2 sound pair is assimilated to an L1 phonological pair, discrimination is predicted to be excellent ('Two Category': TC); in contrast, if the members of a nonnative pair are perceived as pertaining to one and the same L1 phonological category, discrimination is predicted to be poor ('Single Category': SC). A further case is represented by those L2 contrasts whose members tend to be assimilated to a single L1 phonological category, but one fits better while the other is perceived as a bad exemplar of that category ('Category Goodness': CG). In such cases, discrimination is predicted to be rather good, but not as much as in the case of TC; this substantially means that (i) formal instruction will be of great help, and (ii) other – even speech-external – factors will play a role in determining the possibility for a learner to succeed in the discrimination task.

In this sense, the models based on perceived similarity fail to give a satisfactory account of the most critical aspects of perceptual mapping. First of all, as others have suggested, phonological relations need to be viewed as a more nuanced domain than is often assumed in the speech perception literature (Hume & Johnson, 2003). It is difficult to determine what categories like "a varying goodness of fit" and "a good to very good discrimination" actually mean. One possibility for getting rid of this ambiguity, which involves both the phonetic and the psychological levels of analysis, could be to have an L1-oriented look at the problem of L2 correspondences. This might sound paradoxical; yet there are good grounds for attempting to do so.

Consider in particular the phenomenon of allophonic variation, a form of phonetic variation arising from historical as well as synchronic conditions and widely discernible throughout the languages of the world. We unfortunately know little about the psychological status of allophones in speech processing. However, we can appeal to an enormous amount of linguistic facts which are in themselves theoretically very well known: they constitute a rich and highly specified bulk of linguistic data which may be attained for experimental purposes. Could it not be the case that the perceptual discrimination of two L1 allophones has something to do with a more general perceptual process where two members of a speech pair are assimilated to a single phonological category, but one fits better than the other?

If we look at the cross-linguistic problem from such an L1-oriented perspective, we may have at least two advantages. First, we can reduce the problem of the "partially categorizable" speech contrasts to a problem of allophonic contrasts, which patently has taxonomic and heuristic advantages. Second, we will have a methodological improvement, since we will be clarifying some crucial aspects of L1 speech perception, before addressing similar and very controversial aspects of L2 speech perception. This second point would thus be in agreement with G.S. Morrison's view, that unless one has a detailed model of native L1 speech perception, it is impossible to solve the problem of whether an experimental result in L2 speech perception is due to L1 transfer or to some general L2 learning mechanisms independent of the L1 (Morrison, 2006). In particular, following Morrison and colleagues, one should rigorously investigate the native perception of the speech contrast of interest, in terms of native categories, before attempting to understand the perception of the same speech contrast by non-native listeners, in terms of L2 categories.

We believe, then, that the study of L2 speech perception may benefit from research in L1 allophonic processing. This would allow us to have a closer inspection to what exactly happens when a hearer is confronted with two different-in-quality members of one and the same phonological category. As specific purpose of the present study, I will address the question of whether the experience with a neutralization rule in the L1 affects the perception of a corresponding consonant contrast in an L2.

2. The perceptual status of L1 allophonic contrasts

Before addressing the question of the influence of L1 rules of allophonic variation on L2 speech perception, it is necessary to recall briefly the experimental evidence collected on the perceptual status of allophonic contrasts in the native language.

Since the mid-nineties, several experiments have shown that the perception of an allophonic contrast tends to be less accurate than the perception of a phonemic contrast (Pegg & Werker, 1997; Whalen et al., 1997; Peperkamp et al., 2003; Shea & Curtin, 2005; Boomershine et al., in press). In the following paragraphs, I will use the label of 'allophonic effect' to refer to this perceptual phenomenon. The allophonic effect proved true initially for some kinds of context-dependent allophony only, i.e. for those allophones standing in a complementary distribution (see e.g., the alternation between voiceless aspirated and unaspirated stops in English; the alternation between voiced and voiceless uvular trills in pre-consonantal position in French; see also Peperkamp, 2003:102 for some difficult aspects to be handled in the analysis of the complementary distribution).[1] Recently (Hume & Johnson, 2003; Celata, 2006, 2007, in prep.), evidence of an allophonic effect has also been found in cases of 'context-independent' allophony, i.e. for allophones stemming from a neutralization rule and standing in free alternation in a specific context (with one variant generally being preferred over the other). This type of phonological relation is called "partial contrast" by Hume & Johnson (2003). While Hume and Johnson (2003) dealt with suprasegmentals, drawing on perception data on Mandarin tone, Celata (2006 and subsequent) dealt with a consonant contrast similar to those involved in context-dependent allophony; in addition, several experimental paradigms were tested.

The research focused on the Western Tuscan post-sonorant affrication process by which /s/ → [ts] when preceded by /n l r/, both word-internally and in sandhi condition (Stand. It. *orso* → Tusc. [ˈortso] 'bear', Stand. It. *il sole* → Tusc. [il ˈtsole] 'the sun').[2] This

[1] The acquisition of allophonic rules with complementary distribution of variants was also investigated by means of statistical learning algorithms aimed at demonstrating the bottom-up nature of the process of phonological acquisition by infants (Peperkamp et al., 2006, Le Calvez et al., 2007).

[2] As a matter of fact, the process of post-sonorant /s/ affrication is not restricted to some Tuscan dialects; on the contrary, it is widespread in the speech of many central and southern areas of Italy. The strength and modality of application may however vary cross-dialectally.

process is said to have natural articulatory and aerodynamic motivations and may be considered as one instance of the phenomenon of stop epenthesis/emergence within nasal-fricative clusters (Ohala, 1997). The rule in Western Tuscan is in principle exceptionless, but since Standard Italian frequently alternate with Tuscan dialect in the speech of Tuscan speakers (also due to structural and superficial similarities between the two idioms), the [s]-variant may be realized as well, as a minority variant, beside the [ts]-variant, which however is the more frequent realization in that context. From a phonetic point of view, there is complete neutralization between the allophonic [ts] in words like *orso* [ˈortso] 'bear' and the phonological /ts/ in words like *varianza* [vaˈɾjantsa] 'variance' (see Turchi & Gili Fivela, 2004 as far as duration, rise time and energy contour are concerned). These points are summarized in Figure 1. The question mark on the left of [ansa] means that the non-affricated form, as specified above, is a minority variant.

In Celata (2006, 2007), the existence of an allophonic effect for native speakers of a Tuscan dialect was borne out in two identification and discrimination tasks where phonemic labeling and categorical perception (following Gerrits, 2001) were involved, but not when a continuous mode of discrimination was evoked. Three experiments were made, using different experimental paradigms: an AX discrimination experiment, an AXB discrimination experiment, and a 2AFC identification experiment with gating paradigm (AX and AXB discrimination also differed for ISI, which was longer in ABX with respect to AX).[3] Such a complex procedure responded to the intention of testing the effective nature and strength of the presumed allophonic effect under different perceptual circumstances. Although the degree of phonetic (dis)similarity between native and nonnative sounds is the main predictor of perceptual difficulty, it has been shown that several task design variables actually affect the learners' performance (see e.g., Beddor & Gottfried, 1995; Gerrits & Schouten, 2004; Mora, 2005, 2007). Among them, task type, lexical status of the stimuli and ISI appear to determine to a large extent the success of the learners' performance in categorical perception tasks. Exploring several design paradigms in testing the nature of allophonic perception proved useful indeed, since different results were found in the AX discrimination experiment (where allophones appeared to be as easily discriminated as phonemes), with respect to the ABX discrimination and the 2AFC + gating identification experiment (where an allophonic effect was consistently found). This result was interpreted as a consequence of the fact that different tasks evoke different mechanisms of short-term memory. The allophonic effect surfaced only if some degree of categorical perception was requested by the experimental task, i.e. when subjects could rely on pre-existing mental representations of sounds. On the contrary, it failed to emerge in tasks which allowed listeners to compare the stimuli in the auditory sensory memory.

Phonetic Form		Phonological Representation
?[ansa]	→	/ansa/
[antsa]	→	/antsa/
	→	/ansa/

Figure 1. Phonetic outputs and phonological correspondences for the Western Tuscan post-sonorant affrication rule.

[3] The results of the ABX discrimination experiment have been pre-published in Celata (2006); those of the AX discrimination (in comparison with ABX) have been presented at PaPI 2007 (Celata, 2007). A complete account of the whole research on allophonic perception by native listeners, including the results of the 2AFC + gating experiment, is still in preparation (Celata, in prep.). The major results are briefly reported on here, limited to what will be useful for the purpose of the present study.

To sum up, the perception of allophones in neutralization processes appeared to be affected by an allophonic effect, similar to the effect found for allophones in a complementary distribution. This finding is relevant, since allophones in a neutralization context are phonemes in other phonotactic contexts, and their perceptual discrimination is nevertheless less accurate than phonemic discrimination. We can then provisionally conclude that a strictly function-related effect is at stake.

3. Experiment: The perception of nonnative post-sonorant [s]/[ts] contrast by Tuscan speakers

In this experiment, the discrimination of [ns]/[nts] and [rs]/[rts] clusters in a nonnative language by the same pool of Tuscan subjects mentioned in Section 2 is tested. The nonnative language involved is Russian. Subjects were totally inexpert in that language. The hypothesis is that the experience with the neutralization rule in the native idiom (by which /s/ → [ts] when preceded by /n l r/) may influence the discrimination of similar consonant clusters when subjects are perceiving the speech of a foreign language. An experiment of two-alternative forced choice identification task (2AFC) mixed with a procedure of word recognition in noise (Lane, 1963) was submitted to 50 Tuscan and Northern Italian control subjects.

The 2AFC paradigm was chosen because the explicit reference to phoneme categories in the instructions is assumed to promote a phoneme labelling strategy; therefore, perception is fairly categorical (Gerrits, 2001:30). However, the gating technique used in the previous experiment for native allophonic perception (see Section 2) did not prove duplicable for nonnative perception. In fact, when stimuli are presented as isolated fragments, using either Italian or Russian should not produce different reactions in subjects' performance. By contrast, word recognition in noise allowed us to provide subjects with full word (or nonword) stimuli and, moreover, stimuli were encapsulated in full sentences (or nonsense sentences), which provided subjects with a specific linguistic background (either native or nonnative). Half of the stimuli were presented in quiet, half in noise. Several studies have shown that the effect of noise on the intelligibility of foreign-accented speech is somewhat greater than that on native-produced speech (Munro, 1998; Cutler et al., 2007).[4] It is assumed therefore that, if an allophonic effect shapes the perceptual discrimination of nonnative sounds or sound sequences, this would emerge more salient in a case of degraded input.

3.1. Subjects

Twenty-six Tuscan speakers (Livorno and Pisa varieties) and 26 Northern Italian speakers (Piemonte, Lombardia, Veneto, Friuli varieties) were recruited as subjects for the experiment among the students of Università di Pisa. They were paid for collaboration. Since Northern Italian speech preserves the phonemic opposition between /s/ and /ts/ in both intervocalic and post-sonorant position, Northern Italian speakers were used as control group. Henceforth, I will use TUS and NOR to designate the two groups of subjects. All subjects were absolutely inexperienced in Russian.

3.2 Materials

Twelve Russian words and 12 Italian pseudowords in frame sentences were used as experimental stimuli.

---

[4] Different levels of both energetic and informational masking may exhibit different effects on perceptual accuracy. However, masking conditions which cause most misperceptions for native listeners are also those which cause the additional nonnative disadvantge (see e.g., Garcia Lecumberri & Cooke 2006).

As for Russian, the frame sentence was *Ya tebe' povtorja'ju … čëtko* 'I repeat for you… distinctly'. In target position there was one out of 12 words or word sequences containing the clusters [ns], [nts], [rs], [rts], [s:], [t:s] (experimental clusters) and [nt], [nd], [rt], [rd], [t:], [d:] (control clusters). The full list of words is reproduced in the Appendix. The experimental clusters included two post-sonorant allophonic contrasts ([ns] vs. [nts], [rs] vs. [rts]) and one intervocalic phonemic contrast ([s:] vs. [t:s]), for the TUS group. However, as noted above, Northern Italian speech does not exhibit the allophonic rule of post-sonorant affrication, then both post-sonorant and intervocalic contrasts were phonemic for the NOR group. The control clusters were similar to the experimental ones, but included two consonants (/t/ and /d/) which are phonemically distinct in all positions for each group of subjects. Therefore, both post-sonorant ([nt] vs. [nd], [rt] vs. [rd]) and intervocalic ([t:], [d:]) contrasts were phonemic for both TUS and NOR groups.

As for Italian, the frame sentence was the non-sense syllable sequence *Rapotip…latufi*. In target position there was one of the 12 vowel-initial pseudo-words containing the experimental and control clusters cited above for Russian. For the list of Italian materials, see the Appendix. Since previous research has massively shown that listeners make use of lexical knowledge in the phonological categorization of sounds (see e.g., Samuel, 2001; Norris et al., 2003), I decided to use nonwords and nonsense sentences as native (Italian) stimuli. In fact, the experiment was aimed at comparing the perceptual performance for native and nonnative sound pairs with all other things being equal; consequently, possible lexical biases were avoided by using nonwords stimuli for native discrimination and by submitting the task to subjects inexperienced in Russian.

The 12 Russian sentences were uttered by a native Russian female speaker while the 12 Italian sentences were uttered by a native Italian female speaker, recorded in a sound-attenuated room and digitized at 22kHz using an Edirol 4-Channel Portable Recorder and Wave Editor 4-R.

For every sentence, the total duration, the duration of the consonant cluster and the mean intensity were measured. Afterwards, intensity was scaled in order to have uniform values across the sentences (70dB intensity, 0.063Pa RMS amplitude). A silence phase was added at the beginning and the end of every sentence in order to produce an overall duration of 4 sec for the sentence stimuli. These were the sentence stimuli in quiet. In addition, an equal number of sentence stimuli in noise was also prepared, as follows. A segment of babble noise with corresponding duration and intensity values was extracted from the Signal Processing Information Base (SPIB: http://spib.rice.edu/) (4 sec, 70dB, 0.063Pa). The sentence stimuli and the babble noise fragment were then combined additively to form a stereo file which was finally converted into a mono file. The SNR in the overlapped region was a rather adverse one (0 dB).

3.3 Procedure

The perception test employed a two-alternative forced choice identification task (2AFC), mixed with a procedure of word recognition in noise and elicitation of confidence judgments (see Kabak & Maniwa, 2007 for a similar procedure). Subjects were first exposed to the Italian sentences, then to the Russian ones. The procedure was the same in both cases.

Subjects looked at a computer screen where two written alternative (e.g., ANSU and ANZU) appeared, written in capitals in the centre of the screen. Immediately after visual presentation of the stimuli, a sentence stimulus containing one of the two (non)word stimuli was binaurally presented on headphones. Subjects had to decide which stimulus was contained in the sentence they heard, and press the corresponding button on the keyboard. After that, the numbers from 1 to 9 appeared on the computer screen, and subjects had to press the key corresponding to the estimated confidence level of their response: 1 for totally uncertain responses, 9 for totally confident responses, 2 to 8 for intermediate confidence levels. Half of the stimuli were presented in quiet, half in noise (see above). There were no time limits for response.

Subjects performed an initial identification task as training phase (4 sentence stimuli, containing two /p:/ - /b:/ contrasting nonwords), in order to familiarize with the experimental procedure. Subjects were divided in two completely balanced lists.

3.4 Analysis

The number of errors and the mean confidence judgments were analysed. In the error rate analysis, the dependent variable ERROR (correct vs. incorrect response) was crossed with the three-level variable CONSONANT (nC vs. rC vs. VC) in order to test the significance of the main effect. The other relevant variables were NOISE (noise vs. quiet condition), GROUP (NOR vs. TUS), CONDITION (experimental vs. control), LANGUAGE (native vs. nonnative), LIST (list 1 vs. list 2).

Crosstabulations, Pearson's χ^2 (asymptotic 2-sided) and Fisher's Exact test for cells with expected count less than 5 (1-sided and 2-sided tests) were run.

The confidence judgments were initially treated as interval variables but they did not pass the Levene's test for the homogeneity of variance across groups. Therefore an ordinal level of measurement was chosen and the non-parametric Kruskal-Wallis one-way analysis of variance by ranks was run. The significance of the factors CONTEXT (intervocalic vs. post-sonorant) and CONSONANT (nC vs. rC vs. VC) was investigated with respect to the dependent variable JUDGEMENT.

4. Results

Preliminarly to the analysis of the main effect, it appeared necessary to verify whether the general rationale underlying the experimental conditions was respected in the data, i.e. whether the masking condition had an impact on discrimination.

The NOISE factor was significantly correlated with the error rate, both in general (χ^2 (1) = 10.929, p <.01) and on the Italian and Russian stimuli separately (χ^2 (1) = 8.201, p <.01 and χ^2 (1) = 5.431, p <.05, respectively). Moreover, the NOISE factor was also significant in the analysis of the confidence judgments (χ^2 (1) = 135.140, p < .01), with higher scores assigned to the stimuli in quiet than to those in noise condition. Therefore, the prediction about the additional difficulty provided by noise in word recognition was fully supported.

Non-significant factors, on the whole set of data, were LIST, GROUP and CONDITION, while LANGUAGE was significant, with the Russian stimuli bearing, in general, more errors (χ^2 (1) = 16.081, p < .01) and lower confidence scores (χ^2 (1) = 135.039, p < .01) than the Italian stimuli.

The main effect was then investigated. As far as error rates are concerned (CONSONANT * ERROR), the control condition was analyzed first. No comparison resulted significant within this subset of data, nor within the lower levels of analysis (noise vs. quiet condition; for noise, native vs. nonnative stimuli, and TUS vs. NOR speakers). The next step was then the analysis of the main effect in the experimental condition.

Due to the noise effect found earlier, the only subset of stimuli in noise was analyzed. For NOR, no significant effect of the consonant contrast on the error rate was found. For TUS, on the contrary, the effect was strongly significant (χ^2 (2) = 10.438, p < .01). The graphs in Figure 2 show (in percentage) that the error rate for the post-sonorant contexts is significantly higher than for the intervocalic context, as far as TUS subjects are concerned.

The allophonic effect was therefore found in the general performance of TUS, as compared with that of NOR. A further test was however necessary, in order to clarify whether it was the native or the nonnative language or both which bore the effect, and also to investigate whether, for the NOR, a similar effect might be found in at least one of the two cases (recall that the two levels of LANGUAGE are completely independent).

For the TUS group of subjects (see Figure 3), the allophonic effect was clearly borne out for discrimination of the nonnative stimuli (χ^2 (2) = 9.102, p < .05), while it only approached significance for the native ones (χ^2 (2) = 5.275, p = .062). On the other hand, no effect was found for NOR, either on the Italian or on the Russian stimuli.

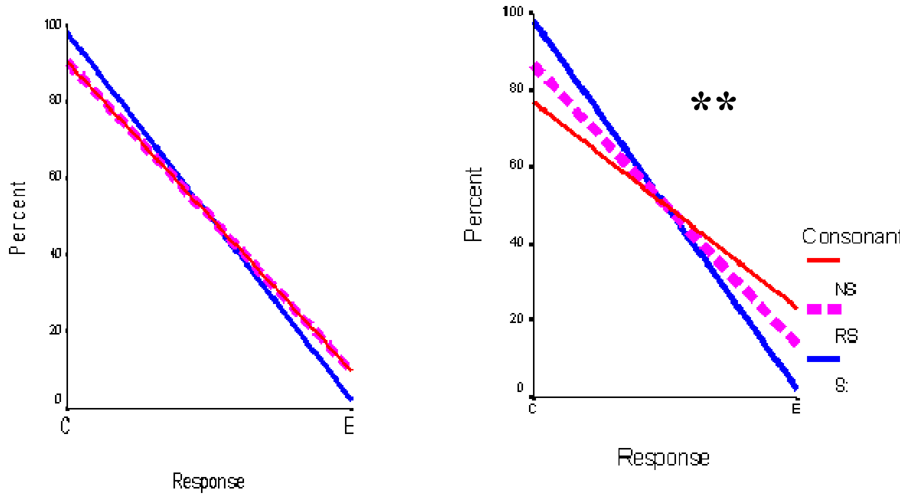

Figure 2. Main effect (CONSONANT*ERROR) for the experimental noisy condition, separately for NOR (left) and TUS (right).

As is evident from the graph on the right in Figure 3, however, the effect burdening on the /nC/ context has a different magnitude with respect to that burdening on the /rC/ context. In other words, some imbalance has to be taken into account between the post-nasal and the post-trill stimuli, where the allophonic effect is concerned. Orthogonal tests were therefore run in order to clarify this point.

For the native stimuli, both comparisons (nC vs. VC, and rC vs. VC) turned out to be significant (χ^2 (1) = 5.879, p < .05 for the post-nasal context; χ^2 (1) = 5.532, p < .05 for the post-trill context), thus suggesting that in this specific case the allophonic effect, even if weaker on the whole, is carried out uniformly by the /nC/ and the /rC/ context. On the contrary, in the case of the non-native stimuli, the /rC/ context was not significantly worse discriminated with respect to the intervocalic context: the allophonic effect appears thus to be exclusively carried out by the /nC/ context (χ^2 (1) = 6.584, p < .05).

One could conclude from this imbalance, that the allophonic effect is in fact a purely acoustic effect, by which different post-sonorant contexts are differently discriminated because of the lower saliency of the target consonant with respect to the intervocalic contexts. However, this should not be the case, because of the following reason. If an acoustic effect was at stake, we should reasonably have found a similar result in the control condition, i.e. for the /t/ - /d/ contrast; yet we didn't. On the contrary, it is worth noting that the different saliency of the post-nasal vs. the post-rhotic context surfaces precisely in the discrimination of nonnative stimuli, and not in the discrimination of the native ones. Therefore, one can conclude that, even if the allophonic effect appears to influence the discrimination of both native and nonnative consonant contrasts (see Figure 2 right), in the perception of nonnative speech some phonetic factors must play some role as well, affecting the discrimination of different types of consonant contrasts in different ways. This point will be further elaborated in Section 5.

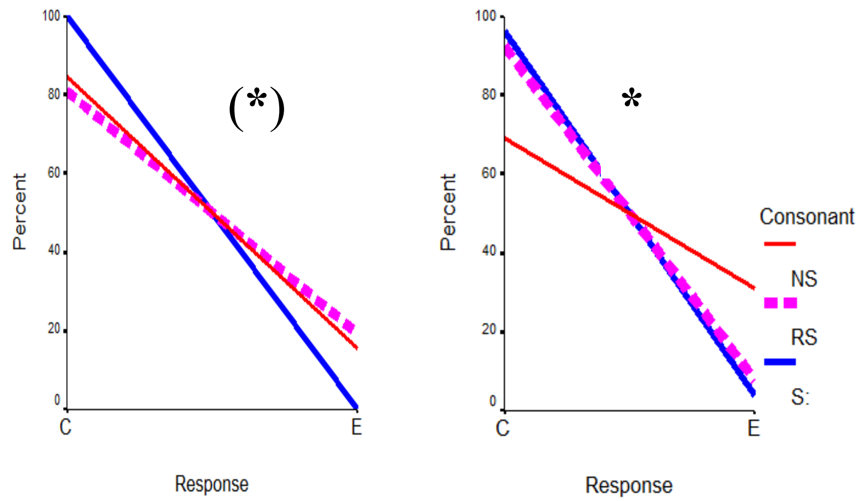

Figure 3. Main effect (CONSONANT*ERROR) for the experimental noisy condition, TUS group of subjects, separately for the native (left) and the nonnative stimuli (right).

As far as the confidence judgments are concerned (CONSONANT* JUDGMENT), no allophonic effect was found either for the TUS, or for the NOR, when the two languages of the stimuli were considered together. But when only the Italian stimuli were considered, the effect approached significance for the TUS group of subjects (χ^2 (2) = 5.837, p = .054).

To sum up, a significant allophonic effect was found out for the discrimination of both native and nonnative post-sonorant stimuli by Western Tuscan subjects. The effect was not always equally strong and appeared to be sometimes conditioned by the interplay of acoustic factors such as the type of the preconsonantal sonorant, but it was altogether present in the performance of Tuscan as opposed to non-Tuscan subjects (who did not show any additional difficulty in the perceptual discrimination of post-sonorant contexts).

5. Conclusions

Two points deserve a final comment, the first regarding the perception of allophones stemming from neutralization processes, the other regarding the influence of native allophonic processes on the discrimination of similar sound contrasts in a nonnative language.

As far as the first point is concerned, the presence of an allophonic effect was borne out for the perceptual processing of Tuscan speakers in an identification task with noise masking. In this respect, neutralizations appear to behave exactly like other allophonic processes, whose variants are in complementary distribution. This result may be supported at least when a categorical mode of perception is involved (see above, Section 2). This means that, for native listeners, two members of a speech pair which pertain to a single phonological category but differ in their goodness-of-fit, may successfully be discriminated when the experimental task allows reliance on pre-existing mental representations of sounds.

As to the second point, the allophonic effect appears to influence in some way the discrimination of *nonnative* similar consonant contrasts as well. However, the effect is not a purely functional one, since some phonetic factors are supposed to play some role as well, affecting the discrimination of different types of consonant contrasts in different ways

This gap is precisely what shapes the non-native perception of partially categorizable sound contrasts, differentiating it from native perception of allophonic contrasts.

Acknowledgments

This research was supported by Italian Governement funding PRIN05/07 (UR Scuola Normale Superiore, Pisa). Technical assistance was kindly provided by C. Bertini & I. Ricci at the Laboratorio di Linguistica of Scuola Normale Superiore, Pisa. Thank you to A. Lentovskaya for reading aloud the Russian stimuli for the test. Useful comments and suggestions for the research project on perception of allophonic contrasts came, as early as in February 2005, from S. Peperkamp and I. Vendelin, and more recently from P.M. Bertinetto and G. Marotta. Thanks to P. Boersma, S. Hamann & K. Dziubalska-Kołaczyk for useful comments to the oral presentation given at New Sounds 2007.

References

Beddor P. S., & Gottfried, T. L. (1995). Methodological issues in cross-language speech perception research with adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 207-228). Timonium, MD: York Press.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp.171-203). Timonium, MD: York Press.

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

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu clicks discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 345-360.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning. In honor of James Emil Flege* (pp. 13-34). Amsterdam: John Benjamins.

Boomershine, Currie Hall, Hume & Johnson (in press) T*he impact of allophony vs contrast on speech perception.* In P. Avery, B. E. Dresher, & K. Rice (Eds.), *Phonological Contrast*. New York: Mouton de Gruyter.

Celata, C. (2006). Percezione e acquisizione di categorie fonologiche: Una riconsiderazione dell'ipotesi allofonica. *Quaderni del Laboratorio di Linguistica*, *6*. [http://alphalinguistica.sns.it/QLL06.htm]

Celata, C. (2007). Perception and acquisition of phoneme categories: the 'allophonic hypothesis' revised. Poster presented at *Phonetics and Phonology in Iberia (PaPI) 2007*, Braga, June 25-27 2007 [http://linguistica.sns.it/PaginePersonali/Celata.htm]

Celata, C. (in prep.). The impact of contrast neutralizations on the categorical perception of speech.

Cutler A., Cooke, M., Garcia Lecumberri, M. L., & Pasveer, D. (2007). L2 consonant identification in noise: Cross-language comparison. *Proceedings of Interspeech 2007*, Antwerp, Belgium.

Flege, J. E. (1987). The production of 'new' and 'similar' phones in a foreign language. Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*, 47-65.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 229-273). Timonium, MD: York Press.

Garcia Lecumberri M. L., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *Journal of the Acoustical Society of America*, *119*(4), 2445-2454.

Gerrits E. (2001). *The categorisation of speech sounds by adults and children: A study of the categorical perception hypothesis and the development weighting of acoustic speech cues*. Utrecht: LOT Publications.

Gerrits E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, *6*, 363-378.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception. The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, *107*, 2711-2724.

Hume E., & Johnson, K. (2003). The impact of partial phonological contrast on speech perception. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the XV International Congress of Phonetic Sciences (ICPhS) 2003.*Barcelona, Spain.

Kabak, B., & Maniwa, K. (2007). L2 perception of English fricatives in clear and conversational speech: The role of phonemic, phonetic, and acoustic factors. Oral presentation at *XVI International Congress of Phonetic Sciences (ICPhS) 2007*, Saarbrücken, Germany. [abstract online at http://www.icphs2007.de/conference/Papers/1380/1380.pdf]

Lane, H. (1963). Foreign accent and speech distortion. *Journal of the Acoustical Society of America*, *35*(4), 451-453.

Le Calvez R., Peperkamp, S., & Dupoux, E. (2007). Apprentissage bottom-up des phonèmes: Une étude computationnelle. *Mathématiques et Sciences Humaines, 180*, 101-114.

Mora, J. C. (2005). Lexical knowledge effects on the discrimination of non-native phonemic contrasts in words and nonwords by Spanish/Catalan bilingual learners of English. In *Proceedings of the ISCA Workshop on Plasticity in Speech Perception (PSP) 2005*, (pp. 43-46). London, United Kingdom.

Mora, J. C. (2007). Methodological issues in assessing L2 perceptual phonological competence. In *Phonetics Teaching & Learning Conference (PTLC) 2007*, UCL London, United Kingdom. [proceedings available online at http://www.phon.ucl.ac.uk/ptlc/ptlc2007_web_procindex.html]

Morrison G. S. (2006). Methodological issues in L2 perception research and vowel spectra cues in Spanish listeners' perception of word-final /t/ and /d/ in Spanish. In M. Diaz-Campos (Ed.), *Selected Proceedings of the 2nd Conference on laboratory Approaches to Spanish Phonetics and Phonology* (pp. 35-47). Somerville, MA: Cascadilla Proceedings Project.

Munro M. J. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition*, *20*, 139-154.

Munro M. J., & Bohn, O.-S. (2007). The study of second language speech learning: A brief overview. In O.-S. Bohn, & M.J. Munro (Eds.), *Language experience in second language speech learning. In honor of James Emil Flege* (pp. 3-11). Amsterdam: John Benjamins.

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

Ohala J. J. (1997). Emergent stops. *Proceedings of the 4th Seoul International Conference on Linguistics (SICOL)*, 11-15 August 1997.

Pegg, J. E., & Werker, J.F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America*, *102*, 3742-3753.

Peperkamp S. (2003). Phonological acquisition: Recent attainments and new challenges. *Language and Speech*, *46*, 87-113.

Peperkamp S., Pettinato, F. & Dupoux, E. (2003). Allophonic variation and the acquisition of phoneme categories. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development*, vol. II (pp. 650-661). Somerville, MA: Cascadilla Press.

Peperkamp S., Le Calvez, R., Nadal, J.-P., & Dupoux, E. (2006). The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition, 101*, B31-B41.

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

Shea, & Curtin (2005). Learning allophones from the input. Poster presented at the 30th Annual Boston University Conference on Language Development. Boston, MA. [also available online at http://128.197.86.186/posters/30/SheaBUCLD2005.pdf]
Turchi L., & Gili Fivela, B. (2004). L'affricazione di /s/ postconsonantico nella varietà pisana di italiano. In F. Albano Leoni, F. Cutugno, M. Pettorino, & R. Savy (Eds.), *Il parlato italiano*, *Atti del Convegno nazionale di Napoli* (pp. 1-25). CD multimediale, Napoli, D'Auria Editore.
Whalen, D. H., Best, C. T., & Irwin, J. R. (1997). Lexical effects in the perception and production of American English /p/ allophones. *Journal of Phonetics*, *25*, 501-528.

Appendix

Russian words

Words	Phonetic trascription	Sentence duration (sec)	Cluster duration (msec)
obed dam	abːjedːam	2.014	155
banket tam	bankjetːam	2.024	165
ta banda	tabandə	1.863	118
dva banta	dvabantə	1.944	134
tri barda	tribardə	1.915	105
ta parta	tapartə	1.841	121
te litsa	tsjelitːsə	1.803	150
melissa	mjelisːə	1.808	155
livantsa	livantsə	1.791	160
romansa	ramansə	1.781	173
adžartsa	adʒartsə	1.799	160
tri barsa	tribarsə	1.872	160

Italian pseudo-words

Words	Phonetic trascription	Sentence duration (sec)	Cluster duration (msec)
addu	adːu	1.46	110
attu	atːu	1.59	131
andu	andu	1.51	128
antu	antu	1.51	147
ardu	ardu	1.50	154
artu	artu	1.55	175
azzu	atːsu	1.45	144
assu	asːu	1.47	152
anzu	antsu	1.43	160
ansu	ansu	1.42	175
arzu	artsu	1.45	170
arsu	arsu	1.42	165