Methodological Perspectives on Second Language Prosody

Papers from ML2P 2012

---

edited by
Maria Grazia Busà, Antonio Stella

cleup

Prima edizione: novembre 2012

ISBN 978 88 6129 937 5

Coop. Libraria Editrice Università di Padova
Via Belzoni – 118/3 – Padova (Tel. 049/650261)
www.cleup.it

TESTING THE PERCEPTION OF L2 INTONATION

Barbara Gili Fivela

CRIL & Università del Salento, Lecce - Italy
barbara.gili@unisalento.it

ABSTRACT

In the first part, the paper offers a review of models, hypotheses and methods found in the literature on the perception of L2 phonemes and of L1 intonation categories. The goal of the review is to point out which methods and hypotheses can be useful for investigating the perception of intonation in L2 and if they need to be adapted for such purpose. In the second part of the paper, a brief summary is offered of an experiment to check if Italian-L1 subjects detect functional differences in English-L2 in the case such differences are conveyed by pitch accents phonetically, but not functionally, similar to L1 pitch accents. The description of main aspects concerning experimental design and results aims to show that: 1) integrating methods and usual practices found in the literature on perception of L2 phonemes and L1 intonation gives interesting results in the investigation of L2 intonation (as shown by reaction time measurements), although, in some cases, methodological adjustments are needed (e.g., splitting oddity discrimination tests); 2) results are consistent with the predictions of degrees of difficulties in discrimination that depend on the phonetic (and functional) features of L1 and L2 intonational patterns, similarly to what proposed by Best's PAM model for phonemes.

Keywords: L2 intonation, perception, reaction time, prosodic transplantation, methods.

1. INTRODUCTION

It is well known that intonation plays a major role in conveying both paralinguistic and linguistic meanings, as it communicates a wide set of information, ranging from paralinguistic content, such as the speaker attitude and his/her emotional state, to information that may affect the computation of the truth value of a given sentence. In this paper, the attention will be focused on the linguistic functions played by intonation and on the way they are learned and perceived when they are expressed in a language that is not native for the subject.

Intonation is usually recognized to convey linguistic information concerning focus, accentuation, phrasing and modality [47]. For instance, the truth value of an utterance may change depending on the marking of focus. Indeed, as Rooth [67] pointed out, given a sentence as "In Saint Petersburg, officers always escorted bellerinas", there are two possible interpretations, depending on where the focus is placed. If "ballerinas" is focused, the interpretation is that officers escorts only ballerinas, and the utterance is a) false, in the case an officer escorts someone who is not a ballerina, or b) true in the case, for instance, ballerinas are escorted by someone who is not an officer; however, if "officers" is focused, the interpretation is that only officers may escort ballerinas, and the truth values of the utterance are the opposite in comparison to what mentioned above. Indeed, the utterance is a) true, in the case an officer escorts someone who is not a ballerina, or b) false, in the case ballerinas are escorted by someone who is not an officer. Thus, in similar cases intonation plays a clear linguistic role. In the literature on intonation, similar types of linguistic information have been argued to be expressed thanks to categorical (quantal) changes that linguists usually attribute to the presence of different phonological categories, as opposed to paralinguistic information that are rather conveyed by gradient modifications [51].

The existence of intonational units as part of the phonological system of a language, composed by contrasting elements within that language, has been proposed within the authosegmental-metrical framework ([3, 62]). In the past thirty years, the existence of phonological intonation categories has been postulated for a number of languages, in which intonation has been shown to express linguistic functions by means of those categories and their combinations (e.g., pitch accents or combinations of pitch accents and phrase accents that express different focus conditions, such as broad vs narrow-corrective focus in various languages; for an overview see [51]). Thus, intonational categories are usually given a phonological analysis, although they are taken to be different from segmental phonological categories, in that they convey very general

meanings that segmental categories do not convey (e.g., introducing an element to the set of shared knowledge, opposing an element to it, and so on [39, 46, 63]). They also play different functions at the sentence level (both linguistic - such as differentiating statements and questions, corrective and neutral broad focus – and paralinguistic functions – such as conveying different degrees of surprise or the speaker attitude). Importantly, there is no necessary one-to-one correspondence between an intonation event and a function as, for instance, an intonational category may actually convey different functions (e.g., in English a rising intonation may be used in both questions and statements expressing continuity, while statements usually show a falling intonation [51]).

Similarly to other phonological units, intonational categories are implemented by means of a set of phonetic characteristics. These characteristics have often been carefully investigated, to the extent that the main phonetic properties of given categories in specific linguistic systems have been clearly specified in terms of average values and main observed trends of variation (e.g., for pitch accents, variation that depends on the number of syllable available in the context or on their distance from preceding pitch events [17, 30, 37]). Thus, the investigations on phonological units and their phonetic form variation have pointed out that there are factors showing a major impact on the way the same phonological pattern may be implemented (see also the reference to individual strategies [59]). Moreover, investigations showed that even quite strong variations in phonetic form do not necessarily correspond to significant differences in the perception of phonological categories and in their association with specific meanings and functions (see, for instance, the correspondences found in the perception of fundamental frequency plateaux and peaks [16]). In particular, a number of works have investigated the perception of intonation categories in order to reach a deep knowledge on the relation between acoustic characteristics and perceived categories (see 3.1.1).

However, the works mentioned above mostly describe languages spoken by native speakers (L1), whereas we do not know much about the relation of acoustic characteristics and perceived categories in a non-native language, and in particular a second language (L2). Indeed, the existing models of L2 perception account for the perception of phonemes and do not address matters concerning intonational competence (see 2.1). Nevertheless, wondering how intonation categories are perceived when they are produced according to an L2 phonetics and phonology is surely of interest.

This paper addresses two relevant issues regarding this topic. First of all, an attempt is made to understand if the core concepts at the basis of models of perception of L2 phonemes may be useful for trying to think of a future model of perception of L2 intonation (with which we will be able to address the issue regarding the general question mentioned in the previous paragraph). Secondly, the paper offers a discussion on which are the methods that appear to be more appropriate in order to successfully perform investigations on L2 intonation categories. In order to address these two topics, the most well-known models of perception of L2 phonemes are described in section 2, together with some observations on how their core aspects may relate to L2 intonation. In section 3, the methods and measurements usually used in applying the abovementioned models are described briefly, and compared with those usually found in investigation on L1 intonation; then some proposals are made on possible adaptations and integrations for investigations on L2 intonation. Finally, section 4 briefly describes an experiment on the perception of L2 pitch accents by Italian learners of English, designed keeping in mind the considerations developed in sections 2 and 3.

2. MODELS OF L2 PERCEPTION

2.1. Models of the perception of L2 phonemes

As already mentioned, existing models of L2 perception basically refer to the perception of phonemes, that is vocalic and consonantal segments. The most well-known models are Flege's ([25, 26, 28]) Speech Learning Model (SLM), Kuhl's ([48, 49, 50]) Native Language Magnet Model (NLM) and Best's ([5, 6, 7, 4]) Perceptual Assimilation Model (PAM). As for the latter, proposals have been made recently to extend it to account for the perception of suprasegmentals, although works performed so far have only dealt with the perception of tones used in tonal languages, such as Mandarin, rather than with the perception of intonation [71, 72]. A detailed description of these models is out of the scope of this paper. However, their main characteristics are briefly outlined here to check if their underlying assumptions may be useful for studying L2 intonation.

For instance, according to Flege's SLM ([25, 28]), "the learning success will be higher" if an L2 phoneme is perceived as different from a native phoneme. In particular, no new category is

established (category assimilation) if members of an L2 category are identified as members of an L1 category, despite audible differences; in these cases, a "merged" category may be created over time that corresponds to the phonetic properties of the L2 and L1 perceived categories. On the other hand, when a new category is established for an L2 sound (category dissimilation) the new L2 category and the closest L1 category move away from one another in the phonetic space (overshoot, to maintain the phonetic contrast in the phonetic space). Flege's model has been developed and tested in various investigations, such as [26, 28]. However it is still based on the idea that the perception of an L2 category and the creation of a new category for the learners are related to the phonetic similarity of L2 and L1 categories.

Kuhl ([48, 49, 50]) proposed the existence of a *Native Language Magnet* effect (NLM, see also Perceptual Magnet Effect - PME), she developed a *Native Language Neural Commitment* hypothesis (NLNC) for L1 perception (NLM e NLM-extended), and finally used them to hypothesize what happens in L2. Kuhl [48] shows that "the prototypes function like magnets that warp the perceptual space": exemplars close to the prototype are not as discriminated as those close to non-prototypical members. Indeed, during L1 acquisition, the discrimination capability develops as highly dependent on L1 characteristics: computational strategies allow infants to identify the statistical distributions of native vowels and consonants and to compute phonemes as the sum of the *phonetic units* they are exposed to. Neural nets are organized accordingly, so that a sort of *Native Language Neural Commitment* takes place ([49]; see NLM-extended for the role of social factors in affecting infant sensitivity to distributional patterns in input [50]). Even though Kuhl's model doesn't directly account for L2 perception, it predicts that L1 characteristics interfere with L2 learning [44], as they work as a "mental filter", affecting the perception of L2 sounds that do not correspond to L1 sounds ([49]: 832). Therefore, according to the model, an L2 sound phonetically close to an L1 prototype is predicted to be not discriminable, as opposed to a sound far from it.

The assumption at the basis of Best's ([5, 6, 7]) *Perceptual Assimilation Model* (PAM) is that listeners extract from the acoustic signal invariants relative to articulatory gestures [12]. Thus, the listener discrimination capability depends on the articulatory-phonetic similarities/dissimilarities s/he perceives between native and non-native sounds. According to this model, given a non-native sound, three situations are possible: 1) the sound is assimilated to native phonemes, with various degrees of goodness; 2) no categorization is possible, as the sound characteristics belong to a neutral phonetic space in between native categories; 3) the sound is non-linguistic, as it is external to the native phonetic space and no inferences are possible as for its articulatory characteristics. In particular, the model predicts six types of assimilation that have an impact on the quality of discrimination of non-native sounds and that, in PAM-L2 [4], are also related to L2 learning. Indeed, in PAM-L2 it is underlined the importance of learning "to perceive the lexical-functional contrasts between the L2 phones", and it is clarified that only for some types of assimilation and discrimination this is expected to be possible. In fact, PAM-L2 predicts that only some learners, after exposure to L2, will develop a new phonological category.

Specifically, the best discrimination is predicted to be possible in the following cases. 1) Two non-native phones are assimilated to two different native phonemes: this is a Two-Category assimilation, corresponding to a very good-to-excellent discrimination and no acquisition of new categories, if the L2 phones are perceived as good exemplars of L1 categories (as a shift to new categories in similar cases is not probable). 2) One of the two non-native phones is assimilated to a native phoneme, whereas the other non-native phone is not systematically assimilated to any native phoneme: this is an Uncategorised-Categorised assimilation that may correspond to a very good discrimination; as for acquisition of the 'categorized' phone, a new category is not likely to be learned, if the phone is perceived as a good instance of L1; on the other hand, the acquisition of the 'uncategorized' phone depends on both phonetic and phonological characteristics of the set of native phonemes with which the L2 phone is assimilated. A quite good discrimination is also possible if 3) neither L2 phone is assimilated to a native phoneme: given this Non-Assimilable type, the model predicts from good to excellent discrimination, depending on the actual differences between the two phones that, in any case, are perceived as non-speech events. These sounds may either be incorporated in the phonological space as uncategorized or be always considered non-linguistic (e.g., they can be produced, but they are not recognized as speech). Conversely, lower discrimination capability is expected in the case of 4) two non-native phones that are assimilated to one native phoneme, but being qualitatively different. This is a Category-Goodness

assimilation, predicting a medium-to-good discrimination and a possibility of acquisition of a new phonological category only for the more deviant phone. Poor discrimination is predicted when 5) two non-native phones are assimilated to one native phoneme as both good/bad instances of it. In this case, a Single Category assimilation takes place and, apart from poor discrimination, the model predicts no acquisition of a new category, unless listeners can perceive phonetic differences between the native phonemes and, at least, one of the L2 phones. Finally, various levels of discrimination capability are expected if 6) neither L2 phone is systematically assimilated to a specific native phoneme, as they are rather assimilated to a set of native phonemes. This is an Uncategorised-Uncategorised assimilation, for which the model predicts from poor to moderate discrimination capabilities. Differences in learning chances are also expected, depending on the similarity of the two phones to the same or to different sets of L1 native phonemes: if the set is the same, one category can be created, as both phones may "converge into a single new but undifferentiated phonetic and phonological category"; if they are different, and thus they are distant in the L1 system, two categories could be created (then, "it is not only the similarity or dissimilarity of a given L2 phone to the closest individual L1 phonetic category that is crucial to perceptual learning, but its comparative relationships within the interlanguage phonological system").

2.2. Towards a model of perception of L2 intonation

The overview of the main features of the models on L2 perception shows that they all share two key aspects: 1) the perception of an L2 category is bound to the comparison of phonetic properties of an L2 instance with that of L1 categories (or their prototypes) and, at least in PAM, it is also related to the phonological organization of the phonetic space and to the "comparative relationships within the interlanguage phonological system" ([4]: 28); 2) the core of the models is the quality of discrimination, as the discrimination capability is the starting point of L2 perception and learning, and it is indeed what is accurately tested.

These two key aspects obviously have a different impact on our considerations on how to investigate the perception of linguistic functions expressed by L2 intonation. Indeed, on the one hand, similarly to what is done for segmental categories, for intonation it is possible to hypothesized a comparison of (phonetic) properties of L2 instances and that of members of L1 categories (or their prototypes), for instance, based on alignment and scaling characteristics. On the other hand, we know that intonation categories are different from segmental categories, as they are argued to have very general meanings and play specific functions at the sentence level. Therefore, working on L2 intonation may require an extra level of investigation in comparison to what is needed for segmental phonemic categories, or, at least, it may require functional and meaning differences to be taken into account earlier in a possible model and in the investigation process. The idea is that in L2 acquisition different assimilation types regard both the phonetic forms and the meanings and functions, and testing the creation of new phonological categories in the case of intonation requires - since the very first stages of the process – to take into account both of them. Indeed, as there is no necessary one-to-one correspondence between intonation events and functions, reference to very specific functions is needed to obtain a precise view of the assimilation types to specific L1 categories in order to predict discrimination capabilities (e.g., if a rising intonation in English is investigated, referring just to 'question' and 'statement', with no further specification, would probably be not sufficient to avoid ambiguity, as a rising pitch can potentially be interpreted both as a question and as a statement expressing continuity).

3. METHODS AND MEASUREMENTS IN INVESTIGATING L1/L2 CATEGORIES

3.1. On L2 phonemes and L1 intonation

The existence and the perception of categories has been widely studied as for both L2 phonemes, for instance with reference to the models mentioned in section 2.1, and L1 intonation, for instance within the autosegmental-metrical framework mentioned in section 1. The various methods and measurements that have been used can be divided in two sets, in line with what was already proposed for methods used in intonation research [33]: speaker metalinguistic judgments, on the one hand, and speaker response and action taking on the other. The former set includes methods and measurements used in behavioral studies in which subjects are asked to judge stimuli, usually thinking consciously about some of their characteristics. The latter set comprises methods and measurements whose goal is to collect data on subject reactions to stimuli, rather than on his/her actual judgments. Methods grouped in this latter set aim to obtain data on what subjects do after hearing stimuli, either in terms of performing

actions or in terms of information processing; in both cases, the experimenter attention is focused on what subjects do interpreting stimuli, also when they do not necessarily think consciously about their properties. Thus, the main difference between the two sets of methods relies on what the experimenters observe and measure during the task.

3.1.1. Metalinguistic judgments

Apart from few exceptions, methods that consist in collecting judgments are used in investigations on both L2 phonemes and L1 intonation, although sometimes with some differences in their implementation. The main ones, if any, are pointed out in this section. The *semantic difference task* [40], for instance, corresponds to asking for a categorical judgment on the presence of a meaning or of its opposite (e.g., question vs statement). It is often used in classification tests in L1 intonation studies and may be used in investigating L2 phonemic categories too, for instance in order to collect judgments on lexical meanings. On the other hand, a *semantic scaling task* [40], that requires asking for judgments on the extent to which a meaning is conveyed (e.g. rating the degree of surprise [15, 42]), appears to be more suited for investigating intonation paralinguistic meanings and it is indeed used for this purpose in investigations on L1 intonation and in the few studies that address issues concerning paralinguistic meaning in L2 intonation [14].

As for methods that regard linguistic categories, checks on the *perceptual equivalence* - or successful imitation - of two patterns were extensively used in the IPO approach for studying intonation [74]. According to this approach, categories – distinct pitch movements – correspond to stylized versions of pitch movements and are identified by means of speaker's intuitions of perceptual equality: two different pitch movements (fundamental frequency tracks) have to be similar to the extent that one is judged to be a successful imitation of the other. In the L2 phonemes literature, a method that can have some similarities (though not for the stylization part) consists in using *intelligibility tests* for assessing L2 production accuracy. In some works, for instance, native listeners are asked to identify items corresponding to L2 productions, given that the L2 is the listener's native language. The aim is to check if productions of segmental categories in an L2 language are perceived by native speakers of that language as perceptually equivalent to members of their native categories (e.g., [27] who asked English listeners to identify English vowels produced by Italian learners of English; [75]).

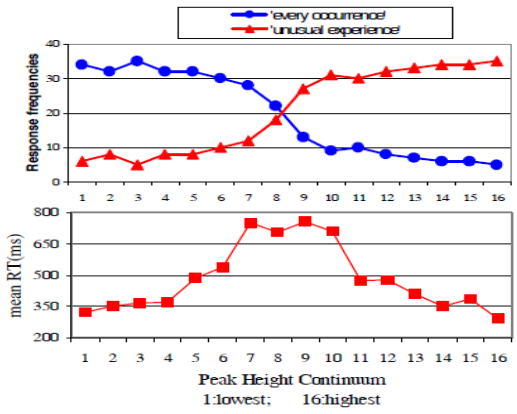
However, the most well-known methods used for investigating categories include identification and discrimination tests. These two tests are part of the *Categorical Perception* (CP) paradigm, originally proposed for testing the categorical perception of consonantal phonemes ([54], for vowels [32]): the combination of their results – in particular the presence of an abrupt change in the identification scores and a corresponding peak in the discrimination scores – is taken to show the existence of two categories that are categorically perceived (the assumption being that linguistic categories are discrete, and members of a category show acoustic and auditory characteristics which are quite stable and perceived as homogenous). Despite the use of the two tasks in investigations on both L1 intonation and L2 phonemic categories, identification and discrimination tests may be used slightly differently in these two scientific fields.

The *identification task* is used to check if stimuli are identified as instances of a category, and is indeed a classification method. In many L1 intonation studies, stimuli are synthetically created so that they represent a continuum of acoustic-phonetic variation between categories, along the lines of the original CP experiments, created on the basis of one or two base stimuli. That is, the continua may be created starting from two base stimuli that are instances of the two categories under investigation (to check for a base effect on the results [76]) produced by the same speaker. Subjects are asked to identify stimuli as members of one of two (or more) intonational categories (for instance, depending on the fact that they can be interpreted as questions or statements), and, finally, data show if stimuli may indeed be divided in two sets and where, along the continuum, is the boundary between those sets (that is, in relation to which stimulus/i and acoustic-phonetic properties the shift from one interpretation to the other takes place, e.g., [45]). Recently, investigators have been using *gating experiments* in order to obtain detailed information on the category shift between intonational phonological units. Indeed, specific classification tests can be used in which, given two base audio stimuli that are instances of two categories, subjects are played segments (gates) of the stimuli of gradually increasing duration, such that the shortest segment may consist of only one/two syllables, while the longest may correspond to the full stimulus [22, 61]. Notice that in L1 intonation studies, data on classification have been obtained also by means of *matching tasks* in which subjects have to identify different

categories, depending on the stimuli interpretation in relation to context information. In such tasks, subjects are given context-stimulus pairs and are asked to judge whether each pair is appropriate or not (e.g., [47] for an overview). On the other hand, to the knowledge of the author, context matching or gating tasks are not used in studies on L2 phonemic competence. Moreover, in various identification tests regarding phonemic competence in L2, stimuli are not part of a synthetic continuum, but rather they are actual members of L2 categories ([27], but see [23] and studies referring, in general, to Kuhl's model, where continua are used for checking the perception of stimuli of increasing acoustic distance from the prototype, e.g. [48]). During classification subjects are asked to identify stimuli as members of one of two (or more) L1 segmental categories. In fact, subjects are expected to assign an L2 stimulus to the perceptually more similar L1 category, even if they do not consider that specific classification as completely satisfactory (goodness-of-fit ratings may then be used to quantify the "level of satisfaction" related to the specific classification – see below). Thus, collected data show which L1 categories are perceptually similar to the L2 stimuli.

In the perception of phonetic categories, subject reaction times (RTs) have also been measured [65], as they offer indications on the cognitive load in decision (thus, they offer hints on the processing of information, similarly to methods and measurements that are discussed in the following section). RTs are often measured during identification tasks run to investigate L1 intonation, as higher RTs point to the presence of a category boundary, offering results consistent with shifts in identification scores [13] – see Figure 1. To the knowledge of the author, RT measures are not used in investigations on L2 phonemes.

Figure 1: Response frequencies and reaction times of the stimuli continuum (adapted from [40]: 99**)**



Apart from the classification of stimuli, or as part of it, there may be an interest in quantifying the quality of stimuli as members of specific categories. In this respect, data may be collected on the goodness-rating/goodness-of-fit of stimuli to the category they are assigned to. This is not part of the traditional CP paradigm, but it may be part of sets of perception tests performed under the assumption that members of a category do not have necessarily to be perceived as homogenous. This is the case of investigations on the Perceptual Magnet Effect (PME - [48]) – see above - where the test is used in order to identify the prototype – best exemplar – of a category. In L1 intonation research on categories, it is not used very often, apart from few investigations on the PME in intonation [70] also in a variant that requires to offer context information ([36, 69]). On the other hand, the test is found very often in the literature on L2 phonemes, as it is used as a measure of the phonetic distance of the L2 stimulus to the L1 category to which is assimilated, or, in any case, as a measure of the perceived relations between categories in two languages ([8, 68]). In some cases, a Fit-Index is also used, combining "both the identification and the goodness-of-fit data into a single metric" (the index is obtained by multiplying the proportion of identification for the goodness rating for that identification [38]: 2716).

As for discrimination tests, they are used to find perceptual thresholds or to check if there is a detectable difference between two sounds. The latter aim is the most commonly found in investigations on both L1 intonation and L2 phonemic competence. Experiments are performed presenting subjects minimum two stimuli and, indeed, experimental designs may vary depending on the number of stimuli compared for each trial, the order of their presentation, and, last but not least, on the question asked in relation to the trial. Differences between tests in L1 intonation and L2 phoneme studies often depend on the type of test chosen. One of the most used discrimination test in investigating L1 intonation is the AX test, involving pairs of stimuli and asking subjects whether the members of each pair are same or different [52]. On the other hand, in the literature on L2 phonemes, discrimination tests are often performed by presenting triplets of stimuli in an ABX or an oddity task. In the former, subjects are asked whether stimulus X is the same as A or B (see, for instance, [43] on differences depending on the order of items in the triplet, e.g., ABX, AXB); in the latter, subjects are asked to indicate the serial position of the odd item in the triplet [38]. Moreover, in the literature on L1 intonation,

stimuli included in discrimination tasks usually belong to a continuum of manipulation between categories (e.g., the continuum used for the identification test in a CP paradigm), obtained from stimuli produced by the same speaker. Besides, stimuli presented for discrimination may be adjacent or not adjacent in the continuum [45], although their reciprocal distance is usually stable (the case of PME experiments is different, as stimuli proposed in pairs are at a variable – progressively increasing - distance from each other, in order to offer a check for the specific hypotheses formulated on the PME existence – [48, 70, 36]). On the other hand, in the literature on L2 phonemes, stimuli included in discrimination tasks are often instances of two different categories (apart from control stimuli), and may also be produced by different speakers [38]. Finally, due to well-documented problems in the application of the discrimination test for studying intonation [55], a variant has been used recently for investigating intonation, offering stimuli that are presented together with context information [31]. The idea behind the proposal is to force subjects to discriminate stimuli with reference to functional information on the difference between them, rather than just on the basis of perception differences that refer to their physical properties.

3.1.2. Processing of information and action taking

Among the methods for collecting data on speaker response and action taking, the most well-known is performing audio recordings by means of the *imitation task*, used in investigating both L2 phonemic and L1 intonation categories. For investigating L2, the method was used for studying the interaction between L1 and L2 phonetic systems [28] and the perception of degree of foreign accent ([24, 26]). In particular, the above mentioned studies apply the *delayed repetition*, introduced by Flege et al. [24] to "prevent direct imitation from sensory memory" ([26]: 162). The imitation task, in this specific variant too, has been used also in investigating L1 intonation and, moreover, for studying categories. The method was proposed by [64] to test whether, listening to a continuum of stimuli that varied for some pitch characteristics, speakers imitated the continuum of variation or they produced instances of two discrete sets. The latter strategy would have meant that they perceived two different categories. The method has been extensively used for intonation research and various modifications have been proposed during the past years (e.g., [41]). For instance, [11] and [33] used a version of the delayed repetition technique in order to favor reference to phonological patterns.

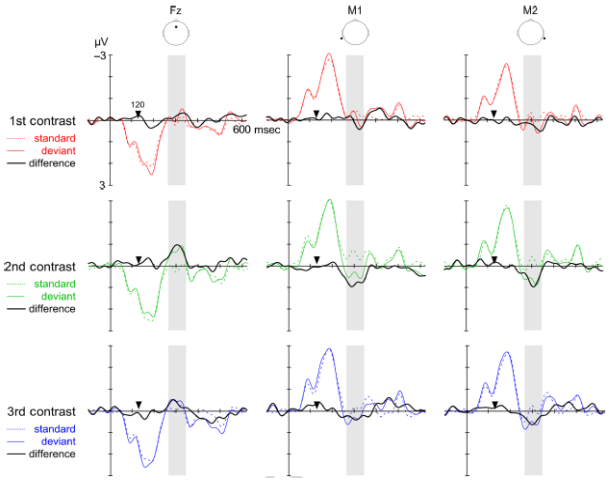
Another method, this time used only in the intonation literature, corresponds to collect data by means of the *card–game task*. This task allows the experimenter to observe the interpretation of stimuli through the acting and the decision taken by subjects during a card-game. The task requires subjects to move two cards depending on audio instructions [33] and allows the experimenter to observe the actions performed by subjects in relation to specific intonation patterns included in the audio instructions. Thus, the collected data may be related to the subject interpretation of audio stimuli.

Within this set of methods, that collect data on subject response and action taking, here we consider also the choice of instruments and measures that allow the experimenter to get data that can shed light on speech processing phases and mechanisms. Apart of the measurement of reaction times - that was discussed above and, in principle, could be included in this section - one of such methods includes the use of *eye tracking*, a technique that allows recording data such as gaze direction and fixation duration. This technology has been quite recently used in investigating both L2 and L1. In the former case, it was used, for instance, to study L2 lexical representations and their phonological content (e.g., within an L2 word recognition task, performed asking Dutch subjects to identify the correct picture, corresponding to competing target words that contained confusable English vowel contrasts for Dutch learners of English [21]). The eye tracking has also been used for investigating L1 intonation categories, for instance by collecting data on the fixation preferences with respect to two figures, as a measure of listeners' interpretation of pitch accent information [18].

Other types of data that relate to the processing of information in speech are collected in neurophysiological and neuroimaging studies, such as those relying on measurements of Event Related Potentials (ERPs), Magneto Encephalography (MEG) or functional Magnetic Resonance Imaging (fMRI). These data have been extensively used in the literature on the L2 phonemic knowledge for finding language-specific perceptual sensitivities, by collecting data on phonemes of different languages and observing the auditory processing of language specific audio information. Indeed, neurophysiological evidence has been first found for the existence of phoneme traces, by recording *ERPs* for the perception of Finnish and Estonian vowels [58]. Various following studies confirmed

the pre-attentive auditory processing of language-specific sounds and confirmed that acoustic contrasts that cross a phonemic boundary elicit specific neurophysiological responses (see, in particular, the Mismatch Negativity component - MMN; e.g., [19, 78]; see [57] for a review). Due to space limits, it is not possible going into details as for tasks and methods that may be used to gather neurophysiological and neuroimaging data. However, it is important to underline that the results mentioned above relate to pre-attentive processing. Such processing is often investigated in neurophysiological studies, by means of encephalography (EEG, for the ERPs mentioned above, that is the recording of brain electrical activity thanks to electrodes attached to a subject's scalp). The recording may be performed while the subject hears audio stimuli and does not pay attention to them. Indeed, for recording data on pre-attentive processing subjects are often asked to watch a silent movie while they are played stimuli. A task that is often used in similar investigations is the *oddball task* in which the stream of stimuli includes an oddball item (deviant) that has distinct characteristics in comparison to other stimuli (standard) and occurs infrequently in comparison to them. Neurophysiological and neuroimaging data have been used less for investigating L1 intonation, at least in relation to the existence of linguistic categories. Indeed, a number of studies related to ERPs data for investigating prosodic and intonational processing and its relations with the syntactic and semantic processing (e.g., [1, 73]), for investigating different global pragmatic meanings in intonational languages, such as in the statement vs question interpretation ([20, 53]). However, the testing of contrasts between intonational categories and, thus, of the existence of phonological representations for specific intonational events is very recent, despite the need and feasibility of similar investigations had been pointed out several years ago ([33, 34]). A first attempt in this direction may be considered [29]'s study on the processing of lexical–tonal and intonational contrasts in an intonational language (Dutch) and a tonal dialectal variety of it (Roermond Dutch). However, evidence for the existence of phonological representations of intonational events is found clearly in [9]'s work, showing that intonational contrasts between statement and question interpretations in Catalan may elicit specific MMN responses; interestingly enough, the results relate to a pre-attentive task similar to that often used for investigating segmental phonemes. In fact, the study supports the hypothesis of phonological representations at the intonational level, as already shown for the phonemic level. In particular, the mean amplitude of the MMN was found to be larger for an across-category contrast in comparison with other contrasts - see Figure 2 (middle panels vs others).

Figure 2: Grand average waveforms elicited to Standard and Deviant stimuli and their difference waves (at three electrodes: Fz, M1, M2). Middle panels represent the across-category contrast (adapted from [9]: 849)



3.2. Toward the choice of the best methods for investigating L2 intonation

There are various differences in the methods used to investigate L2 phonemic and L1 intonation categories, and they relate to instruments, tasks and measurements. Some of these differences are due to: 1) crucial dissimilarities in the objects under investigations, that is, vocalic and consonantal phonemes as opposed to intonational units; 2) the underlying aims of studies on L1 and L2 competence. In the former case, a crucial dissimilarity regards meaning and the fact that segmental categories do not convey it, while intonation categories do. Thus, for instance, in investigating intonation categories there is usually the need to refer to functions or context information to suggest specific interpretations (e.g., can it be used for asking rather than stating information?). On the other hand, as pointed out in 2), dissimilarities may also be due to the presence of different underlying aims in studies on L1 and L2. For instance, L1 studies focus on categories assuming that they should coexist within the same language system and, therefore, one of their aims is often to find out if categories actually exist and are clearly different and separated (to the extent of being discrete) - see the use of continua of stimuli. In L2 studies, the aim is often to find out if L1 and L2 categories coexist in the speaker's phonological knowledge, that possibly includes the two different, L1 and L2, phonological systems; thus,

the aim may be, again, to find out if categories actually exist, but also to check the extent to which they match, they are similar or, eventually, they differ – see goodness ratings.

Thus, the review suggests that methods used in L2 phoneme studies may be adopted in order to investigate L2 intonation, maybe with some adaptations due to the specificities of the object of investigation. In particular, on the one hand, the shift of attention from L1 to L2 intonation easily suggests to adopt the point of view, and related methodological choices, of studies on L2 phonemes (see point 2 above). On the other hand, modifications may be required by the specific properties of intonation in comparison to segmental phonemes (see point 1 above).

For instance, considering the most common experimental design found in behavioral studies on L2, it emerges that a promising design for investigating L2 intonation may include a, say, L2-type of identification test (e.g., stimuli are instances of category members rather than part of continua). Goodness ratings (and fit-index calculation) may be needed for measuring the degree of similarity of L2 and L1 instances, along the lines of what found in the L2 literature. Similarly, discrimination tests performed using oddity tasks appear to be useful to have subjects judging L2 sounds, also produced by different speakers.

However, in thinking of similar methods for investigating L2 intonation, some observations come to mind. The first one regards the use of context information. Indeed, context has been shown to be advantageous for identifying and discriminating stimuli in intonation, due to its efficacy in suggesting specific interpretations and, thus, associations with meanings and functions. Therefore, the use of context information should be kept in mind in experimental designs for investigating L2 intonation.

The second observation regards the need of avoiding the interference from segments. Indeed, an L2 pronunciation of vowels and consonants would represent a possible factor affecting the interpretation of stimuli and their ratings as instances of L1 categories (even though the intonation is at issue). Thus, masking the L2 origin of segmental features in presenting L2 intonational patterns appears to be a desirable characteristic of experimental designs. In this respect, various techniques may be used. For instance, prosodic transplantation allows one to copy the prosodic parameters of a speaker, in this case an L2 speaker, to the productions of another one, here an L1 speaker. Moreover, such technique is often used in L2 studies to investigate the perception of foreign accents (e.g., [10]). Alternatively, delexicalization or altered speech may be used too ([56, 60, 66]), although they seem to interfere too much with the final quality of the stimuli, as they discard information necessary for the message interpretation that is actually required in fine investigations on intonational categories.

Finally, an observation is needed that does not stem from considerations on the specificities related to intonation categories, but rather on the review of methods used in investigating L1 intonation. In particular, it was shown that reaction times are measured as indicators of the cognitive load in decision tasks and offer indications on the processing of information, for instance the processing related to intonation. It seems then plausible that RT measurements could be related to the processing of L2 rather than L1 intonation stimuli as well. Thus, measuring RTs appears to be a promising choice in investigating L2 intonation.

Apart from specific (sub)goals that may explicitly require particular methodological choices, remarks similar to those just mentioned could be made for the other methods discussed in the previous section. For instance, experiments including imitation tasks, as well as measurements techniques as ERPs (e.g., measured in Oddball tasks) or eye tracking (e.g., tracking measured in classification tasks) could easily be used for investigating L2 intonation, although with adjustments such as those mentioned above.

4. PERCEIVING L2 PITCH ACCENTS

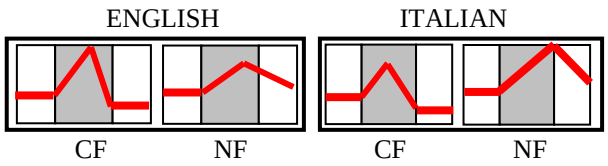
A first check of the usefulness of some of the integrations proposed above can be found in the results of an experiment on the perception of pitch accents categories in English L2 by Italian L1 speakers [35]. The main aspects concerning methods and results are summarized here, to highlight the contribution of the assumptions and methodological integrations suggested in the previous sections.

4.1. An example on Italian L1 and English L2

Aim of the experiment was to check whether Italian (Lecce, Apulia) subjects with a medium-to-low competence and exposure to English as L2 are able to detect variations that signal the non-focal (NF) and correction-focus (CF) pitch accents in English-L2. The NF and CF English pitch accents mainly differ in peak height [2], while in Italian they differ in both height and alignment [77] and, in particular, the CF accent is similar to both NF

and CF English patterns as for alignment and to the English NF for scaling (Figure 3).

Figure 3: Representation of Non-Focal and Correction-Focus pitch accents in English and Italian



Hypotheses on the perception of L2 categories were made with reference to an adaptation of PAM ([5, 6, 7]) to intonation, by considering phonetic properties as alignment and scaling, but also functional properties, that are taken to be a crucial component in the investigation of intonation (see preceding sections). For instance, it was hypothesized that Italians could assimilate both English NF and CF pitch accents to the Italian CF category and could show bad discrimination for these two English accents. However, the association types and the hypotheses concerning discrimination were formulated (and tested) with reference to specific linguistic functions. Indeed, a working hypothesis was that discrimination capabilities could be predicted along the lines of PAM (e.g. good/bad discrimination depends on the types of assimilation to L1 categories) as long as a context was provided for reference to a specific function and interpretation. In fact, due to the lack of a one-to-one correspondence between intonation categories and function, a context 'inducing' a specific interpretation is a way to avoid possible ambiguities that could affect the subject choice during the task.

Three Italian (Lecce, Apulia) and three English (London) speakers for each language produced 5 repetitions of target sentences in which a NF or a CF pitch accent was realized on a target proper name (Guendalina), depending on context information: NF was realized in initial position of neutral, broad focus statements (all new information; e.g. A: Cosa succede*?* - *What's up? B: Guendalina meets Monica*); CF was produced in initial position of correction, narrow focus statements (the correction related to the intial proper name; e.g., A: Maria vede Monica*?- Is Mary meeting Monica? B: N*o, Guendalina vede Monica - *No, Guendalina meets Monica*). A representative production of both the NF and CF interpretation was chosen for each speaker. For each production, the target word was exctracted from the audio signal and manipulated. For each focus condition a resythesis was obtained of three Italian L1 stimuli and three stimuli on which the English prosody was imposed by means of prosodic-transpantation (that is, they showed Italian segments and English prosody). In line with L2 studies – as shown in the previous sections - identification, goodness-rating and discrimination–oddity tests were run. Moreover, as found in L1-intonation literature – see previous sections -, reaction times were measured as indicators of the cognitive load: they were hypothesized to be higher in processing L2 rather than L1 stimuli. Besides, as suggested in L1-intonation literature again, contexts for orienting the intended interpretation were always given (e.g., in the discrimination test too).

Stimuli showing L1 segmental characteristics and L1 intonation and prosody (original resynthesized stimuli – NF-IT and CF-IT) and stimuli showing L1 segmental characteristics and L2 intonation and prosody (NF-IT_ENGprosody, CF-IT_ENGprosody) were played to participants. Eleven Italian L1 listeners (Lecce, Apulia) participated in the experiments.

Results show that, despite a quite high identification score of both L1 and L2 prosody stimuli as instances of the correct category, the latter get lower identification and goodness ratings, in particular, for English prosody NF. Moreover, in line with predictions, subjects have troubles in discriminating English prosody NF and CF.

4.2. Is the methodological integration useful?

4.2.1. Prosodic transplantation

Prosodic transplantation was used to mask the original L2 segmental features, and presenting intonational contour on an L1 base. Other techniques, such as delexicalization or altering speech, were avoided as their impact on the stimuli was considered as not appropriate for the purpose of the experiment. They were taken to interfere too much with the final quality of the stimuli, as they would have discarded lot of information necessary for the message interpretation that was actually required in the linguistic tasks included in the experiment.

Masking segmental features in presenting L2 intonational patterns was considered to be necessary in order to disentangle the impact of intonational-prosodic cues from that of segmental cues. Indeed, an L2 pronunciation of vowels and consonants could have represented a factor affecting the interpretation of stimuli showing L2 target characteristics. Alternatively, original L2 stimuli could also be proposed, similarly to what is done in some works on phonemes in which L2-sounding stimuli are used (e.g., in testing the

perception of English vowels by Italians, vowels are inserted in CVC monosyllabic words that clearly sound English [28]). However, even choosing this latter strategy, a mechanism for masking some of the characteristics of the original stimuli would have been necessary in order to get a homogenous set of stimuli, including L1 and L2 target intonation patterns (indeed Italian L1 intonation patterns were always included among the stimuli, for comparison). In this latter case, segmental characteristics would have had to be masked in presenting Italian target intonation patterns on English sounding base utterances. Thus, the prosodic transplantation technique was considered to be useful in the experimental design and, given the results, appeared to be appropriate for the purpose of the experiment.

4.2.2. Use of context information

Contexts were always provided for suggesting and recalling the intended interpretations (neutral vs correction). Offering context in identification and goodness rating tests was useful, as already shown in the literature on L1 intonation, and had no practical consequences on the design of the experiment on L2 intonation. On the other hand, offering a context in the oddity-discrimination test forced to split the test in two parts, for referring separately to different contexts/functions. Indeed, a short pilot performed asking subjects to point out the odd function, given a triplet of stimuli, appeared to be too complicated. Listening to three productions of the proper name and having to point out which one was odd, meant 1) identifying what function each stimulus played (neutral or correction) and 2) choosing which stimulus was odd. This task was far too complicated for subjects, possibily beacuse it included more than one sub-task and required a too long processing time. The best choice appeared to be to organize two different oddity discrimination tests.

Thus, in one oddity test, subjects were given a context suitable for correction (A: Maria vede Monica*?- Is Mary meeting Monica? B: N*o, Guendalina vede Monica - *No, Guendalina meets Monica*) and, after listening to three productions of the proper name, had to point out the serial position of the item, if any, that could not be used to start a correction (i.e., the odd item in the triplets); in the other oddity test, subjects were given a neutral context (A: Cosa succede*? - What's up? B: Guendalina meets Monica*) and had to point out the serial position of the item, if any, that could not be used to start a neutral, broad focus statement (i.e., the odd item, in the triplets).

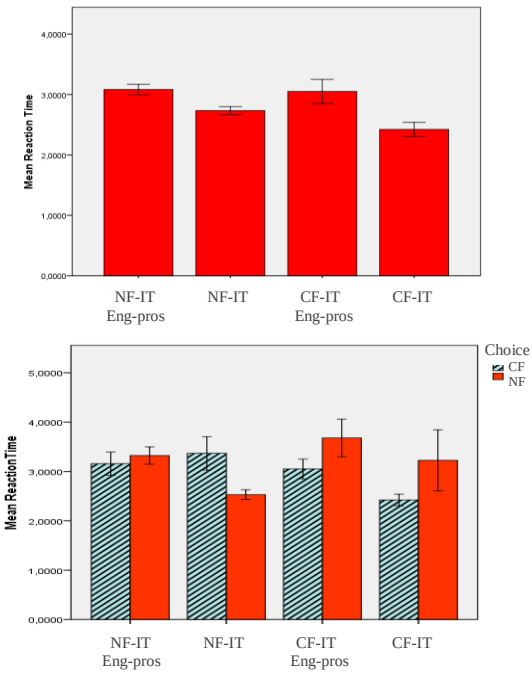
Subjects turned out to be able to perform both oddity tasks, although they showed better discrimination capabilities when they had to point out the presence of a corrective (CF) item out of neutral (NF) items, that is in the latter test described above.

4.2.3. On reaction time measurements

Reaction times were measured as indicators of the cognitive load and were hypothesized to be higher in processing L2 rather than L1 stimuli. Results were consistent with predictions, although they were not always significant; moreover they also showed that, on average, the processing of information was longer in the case of incorrect answers and in the case in which lower identification scores and goodness ratings were found, as well as more difficult discrimination was found due to the type of assimilation of L2 to L1 categories (as it was the case for the assimilation of English NF to Italian CF).

Results showed that RTs (for correct judgments) were longer for stimuli in which L2 prosody was copied than for stimuli with original L1 prosody [ANOVAs: $F(3,1003)=7,331$; $p<.0001$]. In particular, the Tukey post-hoc test showed that RTs were longer for NF stimuli with English prosody in comparison to NF (and CF) stimuli that showed Italian prosody, and they were also longer for CF stimuli with English prosody in comparison to CF stimuli with Italian prosody (Figure 4, upper panel).

Figure 4: Mean RTs for each type of stimulus (upper panel), and for correct/incorrect identification (lower panel, filled/empty colons) (±1 SD)



ANOVAs run separately for data on stimuli NF and CF, showed that RTs were significantly shorter in case of correct identification for CF only (CF: [F(1,3018)=6,082; p<.05], NF: [F(1,302)=2,263; p>.05]). However, RTs were shorter for correct answers to NF produced with L1 prosody, while RTs were alike for both correct and incorrect answers to NF stimuli with L2 prosody (that were probably very difficult to process even in the case of correct identification - see Figure 4, lower panel).

Thus, in the case of our group of Italian L1 speakers with medium-to-low level of competence in English-L2, results on RTs support results on lower identification scores and goodness ratings and suggest a greater cognitive load in the processing of NF stimuli with English prosody.

5. FINAL REMARKS

The paper offered a review of models, hypotheses and methods found in the literature on the perception of L2 phonemes and of L1 intonation categories. The review suggests that working on L2 intonation may require functional and meaning differences to be taken into account earlier in a possible model and in the investigation process. Moreover, the review shows that methods used in L2 phoneme studies may be adopted in order to investigate L2 intonation, maybe with some adaptations due to its specificities.

The description of main aspects concerning experimental design and results of an experiment on the perception of English pitch accents by Italian L1 listeners shows that some method adaptations may include the use of context information and the splitting of oddity discrimination tests; moreover, the use of prosodic transplantation and reaction time measurements may also be useful. Finally, the brief description shows that results are consistent with the predictions of degrees of difficulties in discrimination that depend on the phonetic and functional features of L1 and L2 intonational patterns.

6. ACKNOWLEDGMENTS

I wish to thank Bianca Sisinni for the valuable discussion on the models of perception of L2 phonemes and for her comments on a preliminary version of the paper.

7. REFERENCES

[1] Astesano, C., Besson, M., Alter, K. 2004. Brain potentials during semantic and prosodic processing in French. *Cog.BrainRes* 18, 172-84.

[2] Bartels, C., Kingston, J. 1994. Salient Pitch Cues in the Perception of Contrastive Focus. In P. Bosch et al (eds.), *Focus and Natural Language Processing*, 1-10.

[3] Beckman, M., Pierrehumbert, J.B. 1986. Intonational structure in Japanese and English. In *Phonology Yearbook*, 3, CUP, 255-309.

[4] Best, C., Tayler M. 2007. Nonnative and second-language speech perception. Commonalities and complementarities. In Munro, Bohn (eds) *Second language speech learning.* Amsterdam: J. Benj

[5] Best, C.T. 1994. Learning to perceive the sound pattern of English. In C.Rovee-Collier, L.Lipsitt (Eds.), *Advances in infancy research*, 8, Hillsdale, NJ, Ablex Publishers, 217-304.

[6] Best, C.T. 1994. The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J.Goodman, H. Nusbaum (Eds.), *The development of speech perception*. Cambidge: MIT Press, 167-224.

[7] 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*. Timonium, MD: York Press, 171-204.

[8] Best, C.T., Faber, A., Levitt, A.G. 1996. Perceptual assimilation of non-native vowel contrasts to the American English vowel system, *JASA*, 99.

[9] Borràs-Comes, J., Costa-Faidella, J., Prieto, P., Escera, C. 2012. Specific neural traces for intonational discourse categories as revealed by human evoked potentials. *J.Cog.Neurosc*, 24, 843-53.

[10] Boula deMareüil, P, Vieru-Dimulescu, B. 2006. The contribution of prosody to the perception of foreign accent. *Phonetica* .63,247-67.

[11] Braun, B., Kochanski, G., Grabe, E., Rosner, B.S. 2006. Evidence for attractors in English intonation. In *JASA*, 119, 6, 4006-401l.

[12] Browman, C.P., Goldstein, L. 1992. Articulatory phonology: An overview. *Phonetica*, 49, 155-180.

[13] Chen, A. 2003. Reaction Time as an Indicator of Discrete Intonational Contrasts in English. *EuroSpeech*, 97-100.

[14] Chen, A. 2009. Perception of paralinguistic intonational meaning in a second language. *Language Learning*, 59,2, 367-409.

[15] Chen, A., Gussenhoven, C., Rietveld, T. 2004. Language specificity in perception of paralinguistic intonational meaning. *Lang. and Sp*eech, 47, 311-349.

[16] D'Imperio, M., Gili Fivela, B., Niebhur, O. 2010. Alignment perception of high intonational plateaux in Italian and German. In *Speech Prosody* 2010, Chicago

[17] D'Imperio, M.P. 1999. Tonal structure and pitch targets in Italian focus constituents. In *ICPhS*, San Francisco, 3, 1757-1760.

[18] Dahan, D., Tanenhaus, M.K., Chambers, C.G. 2002. Accent and reference resolution in spoken-language comprehension: Immediate effects of verb based thematic constraints. In *J.Exp. Psyc.Lear, Mem. and Cog.*, 30, 498-513.

[19] Dehaene-Lambertz, G. 1997. Electrophysiological correlates of categorical phoneme perception in adults. *NeuroRep*., 8, 919–924.

[20] Doherty, C. P., West, W. C., Dilley, L. C., Shattuck-Hufnagel, S., Caplan, D. 2004. Question/statement judgments: An fMRI study of intonation processing. *Human Brain Mapping*, 23, 85–98

[21] Escudero, P., Hayes-Harb, R., Mitterer, H. 2008. Novel second-language words and asymmetric lexical access, *J.Phon.* 36,345–60.

[22] Face, T.L. 2005. F0 peak height and the perception of sentence type in Castilian Spanish. *Revista Internacional de Lingüística Iberoam*. 2, 49-65.
[23] Flege, J. 1993. Production and perception of a novel, second-language phonetic contrast. *JASA* 93, 3, 1589-1608.
[24] Flege, J., Munro, M., MacKay, I. 1995. Factors affecting degree of perceived foreign accent in a second language. *JASA*,97,3125–34.
[25] Flege, J.E., 1995. Second-language speech learning: theory, findings, and problems. In Strange, W. (Ed.), *Speech Perception and Linguistic Experience, Issues in Crosslinguistic research*. Timonium, MD: York Press, 233-277.
[26] Flege, J.E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., Tsukada, K. 2006. Degree of foreign accent in English sentences produced by Korean children and adults. *J.Phonetics*, 34,153–75.
[27] Flege, J.E., MacKay, I.R.R., Meador, D. 1999. Native Italian speakers' perception and production. *JASA* 106,5, 2973-2987.
[28] Flege, J.E., Schirru, C., MacKay, I. 2003. Interaction between the native and second language phonetic subsystems. *Speech Communication* 40, 467–491.
[29] Fournier, R., Gussenhoven, C., Jensen, O., Hagoort, P. 2010. Lateralization of tonal and intonational pitch processing: A MEG study. *Brain Res.*, 1328, 79-88.
[30] Frota, S. 2000. *Prosody and Focus in European Portuguese. Phonological phrasing and intonation*, New York: Garland.
[31] Frota, S. in press. A focus intonational morpheme in European Portuguese: Produciton and perception. In P. Prieto, G. Alcibar (eds.) *Prosody and meaning*, Mouton de Gruyter's Trends in Ling.
[32] Fry, D., Abramson, A., Eimas, P., Liberman, A. 1962. The identification and discrimination of synthetic vowels. *Language and Speech 5*, 171-189.
[33] Gili Fivela, B. 2008. *Intonation in Production and Perception: The Case of Pisa Italian*. Edizioni dell'Orso, Alessandria.
[34] Gili Fivela, B. 2009. From production to perception and back: An analysis of two pitch accents. In Fuchs et al. (eds.) *Some aspects of speech and the brain*, Peter Lang, Germany, 363-405.
[35] Gili Fivela, B. in prep. The perception of intonation in L1 and L2: On phonetic variation and function interpretation.
[36] Gili Fivela, B. in press. Meanings, shades of meanings and prototypes of intonational categories. In P. Prieto, G. Alcibar (eds.) *Prosody and meaning*, Mouton de Gruyter's Trends in Ling.
[37] Grice, M. 1995. *The intonation of interrogation in Palermo Italian.* Implications for intonation theory. Tubingen: Max Niemeyer.
[38] Guion, G., Flege, J. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *JASA* 107, 2711-2724.
[39] Gussenhoven, C. 1984. *On the grammar and semantics of sentence accents*. Dordrecht: Foris.
[40] Gussenhoven, C. 2004. *The phonology of tone and intonation*. Cambridge: CUP.
[41] Gussenhoven, C. 2006. Experimental approaches to establishing discreteness of intonational contrasts. In Sudhoff et al.(ed), *Methods in Empirical Prosody Research*, Berlin,de Gruyter,321-34.
[42] Gussenhoven, C., Rietveld, T. 1997. Empirical evidence for the contrast L* and H* in Dutch rising contours. In A. Botinis et al. (ed.), *Intonation: Theory, Models and Applications*, *Proc. ESCA Worksh. Intonation*, Greece, 169-72.
[43] Højena, A., Flege, J.□2006. Early learners' discrimination of second-language vowels, *JASA*,119, 5,3072-3084.
[44] Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., et al. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*, B47-B57.
[45] Kohler, K. 1987. Categorical pitch perception. In *ICPhS*, 331-333.
[46] Kohler, K. 1991. Terminal intonation patterns in single-accent utterances of German: Phonetics, phonology and semantics. In *Arbeitsberichte 25* (AIPUK 25), University of Kiel, 117-185.
[47] Kohler, K. 2006. Paradigms in Experimental Prosodic Analysis: From measurements to function. In *Methods in Empirical Prosody Research*, Sudhoff et al. (eds), Berlin: W. de Gruyter, 123-152.
[48] Kuhl, P. 1991. Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50, 93-107.
[49] Kuhl, P. 2004. Early Language Acquisition: Cracking the Speech Code. *Nature Reviews Neuroscience,5*, 831–843.
[50] Kuhl, P., Conboy, B., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., Nelson, T. 2007. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e), *Philos Trans R Soc Lond B Biol Sci.* 2008, 979–1000.
[51] Ladd, D.R. 1996-2008. *Intonational Phonology*. Cambridge: Cambridge Studies in Linguistics, CUP.
[52] Ladd, D.R., Morton, R. 1997. The perception of intonational emphasis: Continuous or categorical? J. *Phonetics*, 2, 313-342.
[53] Leitman, D., Sehatpour, P., Shpaner, M., Foxe, J., Javitt, D. 2009. Mismatch negativity to tonal contours suggests preattentive perception of prosodic content. *Brain Imag. Behavior*, 3, 284–291.
[54] Liberman, AM., Harris, K-S., Hoffman, H.S. Griffith, B.C. 1957. The discrimination of speech sounds within and across phoneme boundaries. *J. Experimental Psychology* 54, 5: 358-368.
[55] Massaro, D.W. 1998. Categorical perception: Important phenomenon or lasting myth? In *ICSLP*, Australia, 2275-79.
[56] Munro, M.J. 1995. Nonsegmental factors in foreign accent: Ratings of filtered speech. *Stud. in Second Lang. Acq.* 17: 17–34.
[57] Näätänen, R. 2001. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, 38, 1–21.
[58] Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, J., Sinkonen J., and Alho, K. 1997. Language specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432-434.
[59] Niebuhr, O., D'Imperio, M., Gili Fivela, B., Cangemi, F. 2011. Are there "shapers" and "aligners"? Individual differences in signalling pitch accent category. In *ICPhS*, Hong Kong, 120-123.
[60] Pagel, V., Carbonell, N., and Laprie, Y. 1996. A new method for speech delexicalization, and its application to the perception of French prosody. In I*CSLP*, Philadelphia, PA.
[61] Petrone, C. 2008. *From targets to tunes: Nuclear and prenuclear contribution in the identification of intonation*

contours in Italian. Ph.D. Dissertation, LPL, Université de Provence.
[62] Pierrehumbert, J. 1980. *The phonology and phonetics of English intonation*. PhD thesis, MIT – published in 1988 by IULC.
[63] Pierrehumbert, J., Hirschberg, J. 1990. The meaning of intonational contours in the interpretation of discourse. In P.R.Cohen et al. (ed), *Intentions in communication*, MIT, 271-310.
[64] Pierrehumbert, J.B., Steele, A. 1989. Categories of tonal alignment in English. *Phonetica* 46: 181-196.
[65] Pisoni, D.B. Tash, J. 1974. Reaction times to comparisons with and across phonetic categories. *Perc. and Psychoph.* 15, 285-290.
[66] Ramus, F., Mehler, J. 1999. Language identification with suprasegmental cues: A study based on speech resynthesis. *JASA* 105: 512-521.
[67] Rooth, M. 1985. *Association with focus*. Phd Diss., MIT.
[68] Schmidt, A. 1996. Cross-language identification of consonants. Part 1. Korean perception of English, *JASA*, 99, 3201-3211.
[69] Schneider, K., Dogil, G., Möbius B. 2009 German boundary tones show Categorical Perception and perceptual magnet effect when presented in different contexts, *InterSpeech*, Brighton, 2519-2522.
[70] Schneider, K., Lintfert, B., Dogil, G., Möbius B. 2006. Phonetic grounding of prosodic categories. In Sudhoff et al. (eds.), *Methods in Empirical Prosody Research*, Berlin: De Gruyter, 335-362.
[71] So, C.K., Best, C.T. 2010. Cross-Language Perception of Non-native Tonal Contrasts: Effects of Native Phonological and Phonetic Influences. *Language and Speech, 53*, 2, 273-293.
[72] So, C.K., Best, C.T. 2011. Categorizing Mandarin tones into listeners' native prosodic categories: The role of phonetic properties. *Poznań Studies in Contemporary Linguistics 47*, 1, 133-145.
[73] Steinhauer, K., Alter, K., Friederici, D. 1999. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neurosciences* 2, 191-196.
[74] t'Hart, J., Collier R., Cohen A. 1990. *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge, CUP.
[75] Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Hyekyung, S., Flege, J. 2005. A developmental study of English vowel production and perception by native Korean adults and children, *J.Phonetics*, 33, 263–290.
[76] Vanrell, M. 2006. A scaling contrast in Majorcan Catalan interrogatives. *Speech Prosody*, Dresden, Germany.
[77] Vanrell, M., Stella, A., Gili Fivela, B., Prieto, P. under rev. Prosodic manifestations of the Effort Code in Catalan, Italian and Spanish contrastive focus. *J. International Phonetic Association.*
[78] Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csepe, V., Ilmoniemi, R. J., Näätänen R. 1999. Brain responses reveal the learning of foreign language phonemes. *Psychophys*., 36, 638–642.