Anales de Lingüística. Segunda época
n.5 – ago-dic 2020. Mendoza, Argentina ISSN 0325-3597 (impreso) - ISSN 2684-0669 (digital)
http://revistas.uncu.edu.ar/ojs/index.php/analeslinguistica
Recibido: 08.06.2020 | Aceptado: 16.07.2020 | pp. 149-179

Focus on meaning and focus on forms in English intonation categorisation

Foco en el significado y foco en las formas en la categorización de la entonación inglesa

Andrea M. Perticone[1]

Instituto Superior del Profesorado Dr. Joaquín V. González (Argentina)
Universidad del Museo Social Argentino
andreaperticone@gmail.com

Abstract

This study investigated the categorisation of two nuclear pitch accents in English by River Plate Spanish speakers using two different listening modes: listening with a focus on intonational meaning or function and listening with a focus on intonational forms. Thirty-two participants listened to intonational minimal pairs differing in nuclear pitch accents, the forms H* L- L% and H* L- H%. In the first test, they were asked to decide between two pragmatic meanings conveyed by the speaker: making a statement/telling or asking/checking. In the second test, with the same stimuli, they had to decide if the nuclear pitch accent was either falling or falling-rising. The results showed that focus on form promoted a more successful categorisation of H* L- L% as statement, but hindered the categorisation of H* L- H% as question. On the other hand, a focus on forms facilitated the categorisation of the H* L- H% form as falling-rising, whereas it hindered the categorisation of H* L- L% as falling. In view of the L2LP model, it was concluded that a focus on forms promotes more accurate perception of the boundary tone H% within the **new** scenario., whereas it may lead to an erroneous cue weighting in the **similar** scenario. A focus on meaning promotes a more accurate categorisation of forms in the **similar** scenario, but not in the **new** scenario. Thus, a new L2 form becomes more or less readily accessible depending on the listening modality.

[1] https://orcid.org/0000-0002-0127-8106

Keywords: L2 intonation categorisation, focus on form, focus on meaning.

Resumen

Este estudio investigó la categorización de dos acentos tonales nucleares del inglés en hablantes de español rioplatense, según dos modalidades de escucha: con foco en el significado o función entonativa y con foco en la forma entonativa. Treinta y dos participantes escucharon una lista de oraciones dispuestas en pares mínimos entonativos, con igual lexis y sintaxis, pero distintos acentos tónicos, las formas H*L- H% y H* L- L%. En un primer test, los participantes debieron decidir si el hablante estaba afirmando o preguntando; en el segundo test, basado en los mismos estímulos, debieron decidir si el acento tonal nuclear era descendente o descendente-ascendente. Los resultados mostraron que el enfoque en la forma promovió una categorización más exitosa de H* L- L% como declarativa, pero dificultó la interpretación de H* L- H% como pregunta. Por otro lado, un enfoque en las formas facilitó la categorización de la forma H* L- H% como descendente-ascendente, mientras que obstaculizó la categorización de H* L- L% como descendente. A la luz del modelo L2LP, se concluyó que el enfoque en la forma facilita la categorización del tono de frontera H% dentro del **nuevo** escenario, mientras que puede conducir a una ponderación errónea de las pistas prosódicas dentro del escenario **similar**. Un enfoque en el significado promueve una categorización más precisa de las formas en el escenario **similar**, pero no en el **nuevo** escenario. Así, una nueva forma L2 se hace más o menos accesible dependiendo de la modalidad de escucha.

Palabras claves: categorización de la entonación de lengua segunda, foco en la forma, foco en el significado.

Introduction

Intonation, broadly defined as the linguistic use of pitch variations across speech units, comprises three systems: tonality, the division of speech into prosodic units; tonicity, the distribution of accents, and tone, the significant pitch changes located on the tonic or nuclear syllable (Tench 1996:8). The system of English nuclear pitch accents (also known as nuclear tones) comprises five basic phonological variants (Cruttenden 2007:26). These five basic categories are often referred to in terms of their phonetic

forms: falling pitch, rising, falling-rising, rising-falling and level accents. A current debate is related to how fine-grained a description is necessary to distinguish linguistic uses from paralinguistic meanings (Ladd 2006), for example, whether high-pitched and low-pitched falling contours constitute separate phonological categories or two variants of the same phonological category (Cruttenden 2007). There is consensus, however, on two linguistic functions of basic intonational forms: declaratives tend to be associated with falling intonation contours and questionhood is often associated with rising contours.

Within the field of second/foreign language (L2) phonological acquisition, there currently are no models accounting for the processes involved in the acquisition of L2 intonation. However, a number of models of L2 segmental acquisition have been posited which may serve as a starting point for research. These models are based on empirical evidence; they describe processes in the learner's cognitive system and the changes involved in learning L2 speech sounds. The best-known models are the Speech Learning Model (SLM) (Flege 1995), the Perceptual Assimilation Model (PAM) (Best 1995), and more recently—and less widely cited—, the Second Language Linguistic Perception model (L2LP) (Escudero 2005). While these models have epistemologically different bases, they agree on the concept of **perceptual assimilation**: if the learner fails to perceive one or more cues in the phonetic input, a potential distinction will be lost; two different phonological categories will be conflated into a single category within the perceptual space. This may result in an inability to discriminate between two L2 forms or between an L1 and an L2 form.

Intonationally naive learners and even individuals with metalinguistic knowledge of L2 intonation forms and functions may sometimes experience difficulties when categorising intonation patterns. Prosodic features are known to play a role in this process, but there may be extra-linguistic factors as well. This study addresses the issue of how perception and categorisation of two tonal sequences may vary depending on the cognitive mode adopted during listening as well as the prosodic factors present in the acoustic input.

Theoretical framework

According to the Second Language Linguistic Perception model (L2LP) (Escudero 2005), the perceptual categorisation of speech sounds is a language-specific phenomenon. Native perception is the result of an *optimal categorisation* of the L1 sounds. The term 'optimal' is used here in the sense attributed by the Optimality Theory (Prince & Smolensky 1993) and the stochastic Optimality Theory (Boersma 2003).

For second languages, the L2LP model posits that at an initial state of learning, the learners perceive the L2 sounds through an exact copy of the L1 perceptual map. Therefore, learning to perceive L2 sounds in a nativelike fashion implies two cognitive **tasks**: learning which of the auditory dimensions available should be attended to and how they should be weighted. Such **learning tasks**[1] entail the relocation of the learner's perceptual boundaries and the creation of new perceptual spaces, until optimal L2 perception[2] is finally achieved.

The L2LP model proposes three scenarios for the tasks involved in learning L2 sounds: **new** scenario (an L2 contrast not existing in the L1); **similar** scenario (an L2 contrast is similar to an L1 contrast); and **subset** scenario (the L1 has a larger number of categories than the L2, and the L2 categories constitute a subset of the L1 categories). **Learning tasks** will depend on the differences between the initial state (a copy of the L1 perception algorithm) and the final state aimed at, *i.e.* the optimal perception of the L2 categories. In the **new** scenario, the learner perceives fewer sound categories than the ones existing in the L2, and, therefore, an L2 contrast is assimilated into a single L1 category. In this study, it will be assumed that the L2LP model, which is based on experimental research in vowel sound perception, can also be applied to perception of intonation.

[1] Escudero uses the term 'task' to refer to processes taking place in the cognitive system when the learner is exposed to new phonological categories; she does not use the word in the sense of 'classroom activity'. Escudero's learning tasks may be unconscious activities in naturalistic settings.

[2] This is not to say that optimal attainment of nativelike proficiency will inevitably happen; only that it is theoretically possible.

According to the Autosegmental Metrical (AM) approach to intonation (Pierrehumbert 1980; Ladd 1996), the intonation phrase (IP) is made up of a string of abstract tones, high and low (H* and L*), associated to metrically strong syllables. These starred tones are the underlying prosodic primitives of the IP, and surface phonetically as F0 peaks and valleys in the intonation contour. The IP is delimited by a phrase tone (H- or L-) and a boundary tone (H% or L%). The F0 value of an H or L tone is referred to as **F0 scaling**. The location of the H* and L* tones with respect to the strong syllables is referred to as tonal alignment.

The ToBI annotation conventions (Ayers & Beckman 1997), which were developed to represent the AM intonation model, refer to a phonological representation of the F0 contour and not to the F0 contour itself. The H and L tones (starred, phrase and boundary tones) represent the underlying tonal targets within the metrical grid. However, it is common custom in much intonation research to employ the labels H* L- L% and H*L- H% to refer to the actual intonation contours found in the test stimuli, *i.e.* falling and falling-rising pitch accents, respectively[1].

The perception of tonal configuration in an IP consists in the categorisation and weighting of a bundle of temporally integrated suprasegmental features, mainly tonal alignment, tonal scaling, duration, and intensity (Pierrehumbert & Steele 1989; Braun *et al.* 2006). If the starting point in L2 intonation perception is a faithful copy of the L1 system, it is expected that at an early stage, Spanish speakers will attend to parts of the acoustic stimulus which do not have phonological status, and use them erroneously for the categorisation of intonation patterns; on the other hand, it is probable that they will fail to discriminate certain contrasts for which they lack a previously established category.

If the L2LP framework is applied to L2 intonation learning, it follows that there would be a further task involved: learning the form-function mappings valid in the L2, which may or may not match the form-function

[1] For more on the phonetic versus phonological use of ToBi labels, see Arvaniti (2019).

relationships in the L1. In the case of the nuclear pitch accent H* L- L%, this form is also exploited in RP Spanish, as well as the pragmatic function 'making a statement'.

The H* L- H% accent, however, is not part of the Spanish inventory (although a similar form, H*M- has been posited (*cf.* Gabriel *et al.* 2010), at least not as associated to the pragmatic meaning of 'asking a question' or 'checking an assumption' (Brazil 1997), for which River Plate Spanish employs L+H*L-L% (Gabriel *et al.* 2010).

1. Intonation forms and functions

Different tonal configurations can be related to different functions or meanings. According to traditional approaches, a falling intonation contour (H* L- L% in ToBI) can signal syntactic completion or major information (Halliday 1967); the attitudinal approach by O'Connor and Arnold (1973) proposes the labels 'involved, lively, final'; Brazil (1997) proposes that at a transactional level the discoursal function 'proclaiming' signals that the information uttered is assumed to be new to the listener, and therefore lies outside the area of convergence, or common ground, between speaker and hearer.

The H* L- H% sequence, interpreted as a falling-rising contour in the British school, is said to mark dependency relationships between subordinate and main clauses (Wells 2006); in main clauses, it may signal that the speaker implies something which the hearer is expected to know or infer. Brazil (1997) claims that this intonation pattern is used, at a transactional level, to mark information as shared between speaker and hearer. Brazil was the first British author within the small circle of EFL intonation teaching to acknowledge that in modern Standard British English, a referring tone (*i.e.* an H* L- H% pitch accent) on the tonic syllable marks the content of an IP as a *checking question*: the speaker makes an assumption and asks the listener to confirm that it is correct. The tendency to mark checking questions in Standard Southern British English (either in the interrogative or declarative form) has also been noted by Lindsey (2017:91), who claims

that the patterns with L* L- H% proposed by O'Connor and Arnold and Wells are rather dated.

The illocutionary form 'asking' or 'checking' is sometimes marked explicitly by interrogative syntax, but since declarative questions can also have an H* L- H% accent, it turns out that the determination of whether a declarative syntax sentence is being used to state or to ask will depend on either the context in which the IP occurs or the intonation used in the IP. However, if the IP is heard in isolation, the lack of a context will force the listener to rely on intonation alone.

2. Previous studies on the perception of question and statement intonation

It is now acknowledged that intonational meaning is encoded mainly in the nuclear pitch pattern, with the pre-nuclear segment contributing supplementary meaning(s). A seminal study in intonation perception which demonstrated this was carried out by Studdert-Kennedy and Hadding-Koch (1964), from the Haskins Laboratory. They studied the perception and categorisation of English H* L- L% and H* L- H% pitch accents by native American English speakers and Swedish speakers. Subjects were asked to judge utterances both in semantic terms (question or statement) and psychoacoustic criteria (fall or rise at the end). The input consisted in only one phrase, 'for Jane', so that duration remained constant. Different intonation contours were generated by changing F0 scaling at different points: F0 peak, turning point and endpoint. It was found that both native and non-native speakers used both peak F0 scaling and final rising pitch to judge statement or question status in English, and that sometimes high F0 scaling was taken as a cue to questionhood even if there was no final rise. The authors found that there was more intersubject agreement when the first criterion was used. This finding is of special interest for the present study.

More recently, Heeren *et al.* (2015) tested the perception and real time processing of intonation, and their effect on interpretation. They employed

a computer-based game in which the participant played a card game against a computer by means of verbal interaction. In order to test the contribution of intonation to interpretation of meaning, they removed syntactic cues; the stimuli consisted in elliptical utterances of the form "Got a <card category>", which could be elliptical versions of either "I have got a <card category>" or "Have you got a <card category>?". "Got a candy" occurred with a nuclear contour most consistent with a yes-no question (H* H-H%) or a statement (L* L-L%). The placement of the nuclear syllable was kept constant on the last lexical item in the utterance. Real time processing was tested by eye-tracking to assess when and how certain intonational contours affected the interpretation of the utterance. The intended questions versus statements led to different actions on the part of the participant, which in turn correlated with different eye fixation patterns, depending on listeners' online processing of the shape of the nuclear contour, including post-nuclear movement. It was found that the boundary tone (L% or H%) had a decisive effect on the interpretation of the elliptical utterances as questions or statements.

There has been a number of studies related to the acquisition of L2 intonation. Most studies are based on learner speech errors or deviations (Jilka 2000) and the influence of a number of prosodic factors, mainly tonal alignment (Mennen 1998). Many of these studies focus on linguistic perception of intonation contours and their categorisation as 'questions' and 'statements' by L2 learners with different L1 backgrounds. Grabe *et al.* (2005) studied the perception of English sentence intonation by adult English, Spanish and Chinese listeners in order to test the hypothesis that native language influences the perception of similarities and differences among intonation contours. Different versions of the same phrase, carrying different varieties of falling and rising contours, were presented to the subjects, who had to judge the perceptual similarities among them. It was found that all three groups of listeners, regardless of their L1, discriminated the samples with falling pitch patterns from those with rising pitch patterns in the same manner. However, there emerged some perceptual grouping differences among the three groups of subjects. In a second test, subjects

had to judge similarities between sine waves that duplicated the fundamental frequency contours of the speech stimuli. Judgements by English, Spanish, and Chinese subjects yielded no cross-language differences between the perceptual configurations for these nonspeech stimuli.

The consistent cross-language split between falling and rising intonations led to the hypothesis that the perception of intonation contours may start with the activation of universal auditory mechanisms. Two of these allow perception of falling pitch and rising pitch. The cross-language differences led to the hypothesis that the output of universal auditory mechanisms would be shaped by L1 experience, thus yielding different language-specific perceptual configurations.

Zárate-Sández (2018) investigated the perception of high and low boundary tones in English-speaking learners of Spanish at three proficiency levels and compared them with Spanish-English early bilinguals, Spanish monolinguals, and English monolinguals. Categorical perception effects were tested by means of an imitation task.

The results showed that all participants were able to perceive boundary tones in an overall similar manner. Post-test interviews showed that differences in pitch had at least two interpretations: questions (for rising contours) or declaratives (for falling contours). Participants perceived overall pitch height and excursions accurately regardless of proficiency, a perception process which associates the general universal meaning of **open statement** to rising contours and **closed statement** to falling ones (Cruttenden 1981).

As Grabe *et al.* (2005), Zárate-Sández concludes that this rudimentary strategy may constitute the first stage in perceiving and shaping the L2 intonational system, and claims that his findings contradict the hypothesis that this intonational pattern in Spanish would pose a challenge for English speakers. He proposes that perception of Spanish boundary naturally develops from an early stage and approaches native-like processing at high proficiency levels. Thus, minimal to no explicit pronunciation instruction

would be required. He goes on to state that, since the performance of monolingual English speakers was similar to that of bilinguals, it could be posited that accurate perception of L2 boundary tones can be successful even if the hearer does not possess familiarity with the L2.

A common factor in intonation studies is the fact that the subjects recruited for the tests tend to be 'intonationally naive,' *i.e.*, with no previous explicit instruction in the intonation of English as an L2, no systematic training in the identification of intonation patterns nor knowledge of intonation theories. In this study, all subjects had such training and knowledge.

Aims of this study

The aim of this research was to find out whether Spanish speaking subjects with previous instruction in English intonation would categorise two English intonation patterns in different ways, depending on the nature of the listening test and the cognitive modality entailed by each.

It was assumed, in accordance with previous research, that intonational meaning is mostly derived from the nuclear pitch pattern, with the pre-nuclear segment contributing supplementary meaning to the meaning encoded in the nuclear accent.

The tonal sequences investigated were the nuclear pitch accents H* L- L% and H* L- H%. The specific aim was to assess the subject's categorisation of these intonation patterns in a forced choice test involving two specific pragmatic meanings or functions: illocutionary acts of making a statement or **proclaiming** (Brazil, (1997) and acts of asking a **checking** question.

In contrast to previous studies, the H* L- H% nuclear accent was chosen instead of the simpler rising contours H* H- H% and L* L- H%, since the H* L- H% pattern has become widespread in standard Southern British English checking questions (Lindsey 2017), replacing the once preferred form L* L- H% found in many dated English intonation manuals still used in English Teacher Training Colleges in Buenos Aires, where this study was conducted. Indeed, the participants in this study were trained to favour the H* L- H%

form for checking questions, be they interrogative, declarative, or elliptical in form.

More specifically, the focus was placed on a) the distribution of the participants' categorisation of the intonation input, *i.e.* the frequency with which the stimuli with the H* L- L% sequence was categorised as 'statement' and the frequency with which the H* L- H% stimuli were categorised as 'asking/checking,' and b) the comparison of their performances with those in a second test based on the same acoustic stimuli but using a different test, which involved listening and categorising according to psychoacoustic criteria, with a focus on forms, *i.e.* whether the nuclear (tonic) pitch accent was judged to have a falling pitch direction or a falling-rising one. The assumption was that, if participants categorised a given tonal sequence as conveying 'asking/checking' and as 'falling rising,' this was due to their perception of the boundary tone H% in the nuclear pitch accent H* L- H%, which is one of the main differences between the form in the **new** scenario in the L2LP framework and that in the **similar** scenario.

Since the pattern H* L- H% does not occur in River Plate Spanish checking questions, it was expected that participants would be less successful at perceiving H% than L% (or discriminating H% from L%) and would therefore incorrectly associate some of the H* L- H% input tokens with acts of 'stating' and with 'falling.'

A further aim was to examine the relative role of duration in the perception of the H* L- H% in minimal pair words. It was expected that the perception of H* L- H% in minimal pair words would be favoured by the longer duration of the period between the turning point and the endpoint.

The test conditions intended to promote the employment of intonation cues as the sole source of categorisation, since the sentences were decontextualized, as in the studies conducted by Hadding-Koch and Studdert-Kennedy (1964), replicated by Zárate-Sández (2018). Thus, the listeners had to rely entirely on the intonation forms in order to arrive at a categorisation, regardless of the listening mode adopted.

Method

1. Participants

A total of 32 River Plate Spanish speakers participated in this study. Their ages ranged from 19 to 37 years old, with an average of 26.2 years (SD=4.6). They had previously received instruction in the forms of English intonation and their meanings, as well as training in identification of intonation patterns.

The participants' general level of proficiency in English was assessed using the **C-test** (Keijzer 2007; Schmid 2020). Scores ranged from 58 to 98 marks out of 100, with a group mean of 79.5 (SD=9.7), equivalent to C1 (Proficient User, bottom band) within the Common European Framework of Reference.

Only three out of the total number of subjects had had musical training: of these, two had sung in a choir, and one played a musical instrument. All subjects reported normal hearing.

The research project was presented to potential participants as 'research to analyse aspects of English speech' and publicly advertised at Teacher Training Colleges and on social media. No mention of the terms 'intonation' or 'tones' was made in the accompanying flyer. The aim of this omission, as explained in the Procedure section, was to prevent subjects from approaching the tests with the idea that they were going to be tested on the identification of intonation forms, which would have most probably led them to approach Test 1 with a forms-oriented strategy rather than a holistic listening mode necessary to access pragmatic meanings, *i.e.* telling or asking.

Participants were all students at a Teacher Training College in the city of Buenos Aires, Instituto Superior del Profesorado Dr. Joaquín V. González. All subjects had attended both Phonetics and Phonology II and Laboratory Practice II courses, where they received an eight-month instruction in the

forms and functions of English intonation within the British tradition (O'Connor & Arnold 1973; Brazil 1997).

While they had been taught with a **Focus on Form** approach (Long 1991) for the introduction of new form-function mappings, the perceptual and oral training carried out at the early stages was based on **a focus on intonational forms**. This means that participants were already familiar with the task of identifying English **nuclear tones** (pitch accents) by attending solely to their phonetic forms, employing psychoacoustic criteria to relate F0 changes occurring on the nuclear syllable (and on post-nuclear syllables, when applicable) to a pseudospatial scale.

2. Materials

The stimulus material consisted of a set of 18 pairs of target sentences and 8 fillers read by a male native speaker of Standard Southern British English. Each pair differed only in the nuclear pitch accent used on the last lexical item: there were 18 utterances with falling intonation (nuclear tonal sequence H*L-L%) and their 18 counterparts bearing a falling-rising nuclear sequence (H*L-H%).

The H* L- L% versions of the sentence had a mean F0 peak of 150 Hz (corresponding to H*), and the endpoint (L%) had a mean F0 of 86 Hz. The H* L- H% versions had a mean F0 peak of 150 Hz, a mean turning point (L-) of 86 Hz and an endpoint (H%) varying from 110 to 140 Hz, depending on the duration of the syllable, or the presence of a post-tonic syllable.

The form was declarative for both members of each pair, so that listeners would have to rely solely on the prosodic level in order to decide which of the pragmatic meanings, 'statement' and 'asking/checking', was intended by the speaker.

A: They tied the KNOT. (H* L- L%)

B: They tied the KNOT? (H* L- H%)

In all 36 sentences (see Appendix), the nuclear pitch accent was placed on the last lexical item. The items used varied in the voicing duration of the period starting in the starred syllable and ending in the boundary tone. The shortest syllables contained lax vowels in the nucleus and unvoiced consonants in the coda, *e.g*. 'bit,' 'bus,' 'knot'; longer syllables contained lax vowels followed by voiced consonants, *e.g*. 'bid,' 'buzz,' 'nod'; there were also pairs containing tense vowels, as in 'seat,' 'bead,' 'fought,' 'ford'; the longest one-syllable items contained tense vowels and no coda, *e.g*. 'sea,' 'four.' Thirty-one pairs of utterances had no post-tonic syllables, while the remaining five had one post-tonic syllable. In these cases, the minimal pairs varied only in the segmental composition of the post-tonic syllable, thus yielding different durations: 'leave it,' 'leaving'; 'icebox,' 'iceberg.' In this way, different phonetic implementations of H* L- L% and H* L- H% were obtained.

Since the native speaker was asked to sound as natural and colloquial as possible, it was inevitable that the scaling of H* varied slightly from token to token. Therefore, in order to avoid the presence of **high key/termination** (Brazil 1997) and the meanings it projects (**contrast** and **adjudication**, both of which may contribute to a degree of perceived questionhood), the F0 of H* tones were normalised using Praat v. 6.0.47. Likewise, the F0 scaling of H% was normalised whenever possible. In some cases, it was not possible to manipulate the intonation patterns so that they would exactly match one another: under certain conditions, broad manipulations of F0 introduce distortions in the acoustic signal and can create a 'mechanical voice' effect. Since the aim was for the audio tokens to sound as natural as possible, a small degree of accuracy was sacrificed for the sake of authenticity.

A further treatment of the stimuli was aimed at ensuring that the pre-tonic segments in the H* L- L% and the H* L- H% utterances would be identical, so that the intonational contrast would result only from the tones in the tonic and post tonic segments. This procedure precluded another potential cue to questionhood, which is the presence of high pitch throughout the pre-tonic as well as through the tonic. In this way, a further source of

variation was controlled for. **Figure 1** shows the F0 of the relevant tones, in Hz, and the result of the normalisation so that the pair would be almost identical, only differing in their edge tones L-L% and L-H%.

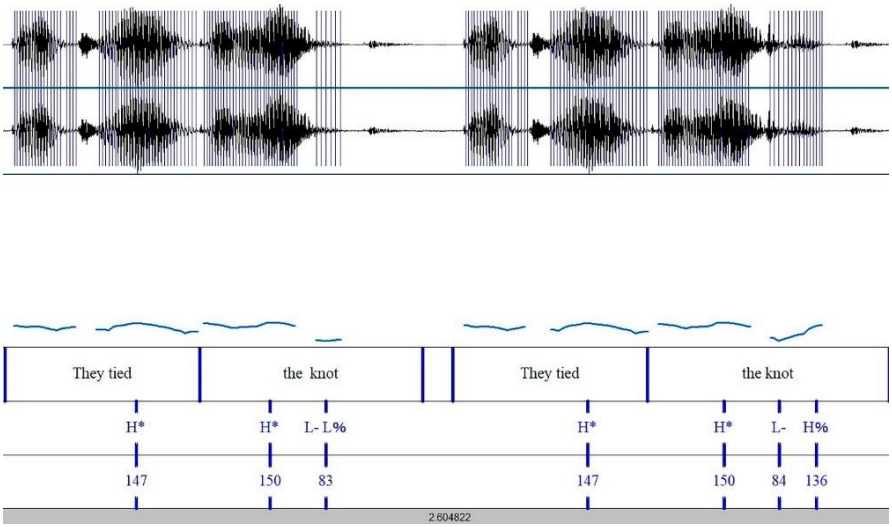

Figure 1. Intonational contours of 'They tied the knot' in the stimulus. The pre-tonic segments are identical. The tokens differ in the nuclear tonal sequence associated with the last lexical item: H*L- L% and H* L- H%. Note the F0 scalings of the starred tones and the phrase tones are virtually identical.

3. Procedure

3.1. Test 1

Test 1 consisted in a listening identification test involving a synthetic, holistic approach, with a focus on pragmatic meaning. Participants were asked to picture two situations: a speaker addressing a question to a listener and a speaker making a statement to a listener. They were told to listen to each sentence globally, intuitively, *i.e.* without paying attention to any specific linguistic feature or analysis, and decide between two possible pragmatic meanings conveyed by the speaker: making a statement/telling something or asking/addressing a checking question to a listener. According to their categorisation, they had to click on 'statement/telling' button or the 'asking/checking' button.

Since falling intonation patterns can also function as questions, although of a different type (finding-out questions), the participants were told beforehand not to consider this option.

Since the participants were all students with metalinguistic knowledge of intonation, the term 'intonation' itself was not used, nor were any 'tones' mentioned. Any reference to the forms of intonation was carefully avoided, so as not to induce an unwanted analytic listening mode which would promote a focus on forms. Participants were not tested during their Phonetics or Laboratory classes. The subjects did not know the purpose of the tests other than 'explore aspects of English spoken language.' The participants were also told that the test sessions were open to anyone with some degree of general English, and that the tests did not require any kind of metalinguistic knowledge (naive participants' results were set aside for further research).

It was expected that all participants would tap into the ability to perceive the L-H% edge tones and associate them to the pragmatic function **question** or 'asking/checking'.

For the test administration, software TP (v. 3.1) was used. For each utterance, participants were instructed to listen and click on either the 'statement' or 'question' button. The test was self-paced and each sound file could be played twice. Prior to the test, participants were given a short rehearsal session.

3.2. Test 2

Test 2 was performed in a separate session. This time, participants were requested to listen to the same utterances as in Test 1, but this time with an analytic, psychoacoustic listening mode. They were told that, in all utterances, the nuclear pitch accent was placed on the last lexical item. They were asked to pay attention to identify the nuclear tone sequence in each utterance. The choice was forced between 'falling' for the nuclear accent H* L- L% and 'falling rising' for H* L- H%.

As in Test 1, for the administration of Test 2, software TP (v. 3.1) was used. This time, participants were instructed to listen and click on either the 'falling' button or falling-rising' button. The test was self-paced and each sound file could be played twice.

Results

1. Data collection and tabulation

The data were collected by the TP software in csv spreadsheets: subject ID, sentence token, subject's responses, number of hits, and reaction times. The latter were set aside for future studies.

Results were tabulated into two sets: responses to 'statement' and 'falling' were grouped together under the label H* L- L%; those for 'question' and 'fall-rise' were grouped under the H* L- H% label.

The C-test scores and participant ages were also recorded in the data set.

As for the coding of responses, for each correct identification of pragmatic meaning each subject was awarded 1 point. If an utterance had a H* L- H% nuclear accent and the subject had clicked on the 'statement' button, no point was awarded. Each subject's final score was the frequency of correct responses, or number of hits, out of 36 utterances heard, of which 18 were H* L- L% and 18 were H* L- H%. The scores were converted to relative frequencies (*i.e.* a scale of 100).

2. Listening test results

Table 1 below shows the distribution of scores. The group hit mean when categorising the H* L- L% tonal sequence correctly as 'statement' was 89.44 (SD=10.25), while the same sequence was judged as 'falling' 81.85% of the times (SD=10.51). The H* L- H% sequence was judged as 'question' 65.37% of the times (SD=17.73) and the mean hit rate was 84.26 for 'falling-rising' (SD=11.31). These results are summarised in **table 1** below.

Table 1. Overall relative frequencies of correct categorisations for each nuclear accent.

Nuclear pitch accent	**H* L- L%**		**H* L- H%**	
Categorised as	**Statement (Test 1)**	**Fall (Test 2)**	**Question (Test 1)**	**Fall-Rise (Test 2)**
Rel. Frequency (%)	**89.44**	**81.85**	**65.37**	**84.26**
SD	**10.25**	**10.51**	**17.73**	**11.31**

Group results show that the participants were more successful at categorising H* L- L% in Test 1 (89.44%) (synthetic listening and pragmatic function) than they were in Test 2 (81.85%) (analytic listening and psychoacoustic representation of the intonational forms). With the exception of three outliers, the group responses for 'statement' show less variability than the responses for 'fall.'

The boxplot in **figure 2** below shows the responses for each category: statement, fall, question, fall rise.

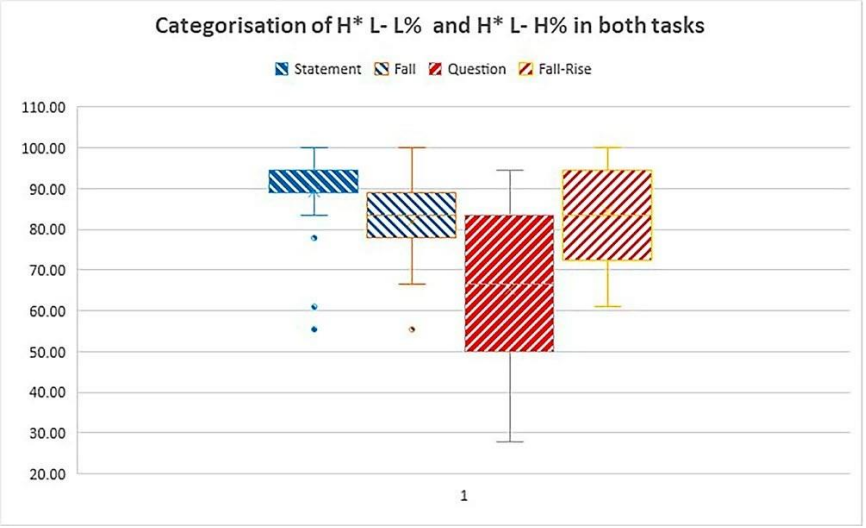

Figure 2. Comparison of results and distribution of scores for each categorisation in each test.

3. Statistical treatment

In order to test if the listening mode and the type of focus employed in the tests had an effect on the categorisation of the H* L- L% and the H* L- H% forms, a Wilcoxon Signed Rank Test was run to check on the cross-test differences for each intonational form. The test showed a significant difference between scores in Test 1 and Test 2 for each nuclear intonation pattern. For the H*L-L% form, Z=-2.342, p=0.019; for H*L-H%, Z=-4.04, p=5.4 E-5 (α=0.05). The categorisation performance for each nuclear pitch accent varied considerably across tests; the greater degree of variability was shown by the H* L- H% accent.

Analysis

The tonal sequence existing in the L1, H* L- L%, was processed and accessed more efficiently when the synthetic mode was used (focus on meaning). This is in accordance with the findings by Hadding-Koch and Studdert-Kennedy (1964). However, for the sequence H* L- H%, which does not occur in the L1, or which does not have a questioning function in the L1, the reverse held true: the new L2 form H* L- H% was processed and accessed more successfully when the analytic mode (focus on forms) was used.

The categorisation of almost 90% of H* L- L% tonal sequences as 'statement' is in line with Escudero's claim that in the **similar** scenario, the learner's task will be relatively easy. The H* L- L% form also occurs in Spanish and has the same function as a prosodic marker of the illocutionary act of stating as it does in English; hence, the high frequency of success.

However, it may seem rather striking that in Test 2 the H* L- L% pitch accent was judged to be 'falling-rising' 18% of the times. A possible source of error in the categorisation of H* L- L% as a falling-rising contour may be incorrect attention to and weighting of prosodic cues. The subjects most

likely heard rising pitch somewhere in the utterance. Rather than the F0 peak for H*, and the F0 endpoint for L-L%, they may have paid attention to parts of the F0 **interpolations** between different points of the utterance. According to the AM Model (Pierrehumbert 1980), these interpolations are phonetic rather than phonological. One interpolation is that of the step-up leading to H*, which may include a part of the pretonic segment (as in 'They'll leave it', 'She's leaving' and 'They hit an iceberg') and /or the pre tonic taken as the only cue, whereas the phrase and boundary accents were discounted. Escudero (2005) explains that in the **new** scenario, the **learning task** consists in learning what parts of the acoustic input constitute cues, which parts should be discounted, and how the cues should be weighted, *i.e.*, which cue functions as a primary one and which one(s) are of secondary order. If the learner pays attention to phonetic information which does not act as a cue, the stimulus will be assigned to the incorrect phonological category.

The categorisation of H* L- H% sequences suggests that participants weighted prosodic cues differently for the same nuclear word. **Figure 2** above shows the F0 contours for the utterance 'They hit an *iceberg′ with H* L- L% and with H* L- H%. The H* L- L% form was judged to have a falling contour less often (23.3% of the time) than it was judged to constitute an act of proclaiming (80%); however, when this same utterance carried an H* L- H% accent, the intonational form 'falling-rising' was detected more successfully than the illocutionary force 'asking a checking question.' It is likely that as the H* L-L% pitch accent is also present in the Spanish inventory of intonational forms, participants were able to perceive the tones, and process the sequence in the same way they do in everyday communication, *i.e.* with a focus on meaning, or synthetic listening mode. On the other hand, since H* L- H% is a new L2 form, regardless of the high boundary tone H%, the subjects failed to associate the form with the meaning 'question,' a mapping nonexistent in Spanish.

It should be borne in mind that, within the L2LP framework, the **learning task** involved in the acquisition of new phonological categories is twofold: the learner has to learn not only the new L2 forms, but also the new form-

function mappings. In Test 1, the **new** scenario included both the H* L- H% tonal sequence as a form, and the H* L- H% question mapping. While the subjects had metalinguistic knowledge of this form-function association, it seems that it remained at the level of declarative knowledge rather than procedural. However, there remains the question of why, if the choice was forced between the categories 'statement' and 'question/asking', they preferred to categorise a H* L- H% contour as a statement. The participants were aware that 'implication' or any other possible meaning of this intonation pattern was not an option. It follows, then, that the most likely reason why they did this is that they failed to perceive the H% (or did not weight it as a primary prosodic cue) and selected the H*L- interpolation as the only cue, thus assimilating the contour to the already known one, the falling contour.

The results showed that the perception and categorisation of the intonational sequences H*L-L% (falling pitch) and H*L-H% (falling-rising pitch) may depend on the type of listening mode employed. It is probable that in the **similar** scenario, a focus on function promotes a more successful categorisation of H* L- L% sequences as statements, but hinders the categorisation of H* L- H% sequences as questions; since the H* L- H% question form-function mapping is a new element, it will be more difficult to learn. On the other hand, a focus on forms might facilitate the categorisation of the H* L- H% form as falling-rising, whereas it could hinder the categorisation of H* L- L% as falling. This, in turn, would mean that a focus on forms, with its reliance on perceptual attention strategies, promotes the perception of the boundary tone H%, whereas this type of focus places limitations on an accurate perception of L%. In the light of the L2LP model, it could be concluded that a form existent in the L1 becomes more or less accessible depending on the listening modality.

There arises the question of how the degree of cognitive complexity may promote or prevent successful categorisation. The construct of **task complexity** (Robinson 2005) has been devised with production tasks[1] in

[1] In this case, the term 'task' has been used in the sense 'conscious activity'.

mind, mainly speaking and writing. In the context of this study, it could only be applied to the task-oriented activity performed in Test 1, listening for pragmatic meaning. Despite differences between a strict interpretation of the term 'task' (Ellis 2003) and a generic use of the term 'task' employed here, it may be useful to compare the levels of complexity in Tests 1 and 2 in terms of the demands of linguistic processing imposed by each.

Whether the synthetic listening mode employed in Test 1 is more or less complex than the analytic mode in Test 2 will depend on what task features are taken into account. If familiarity with the task is taken to mean 'less complex,' it is clear that Test 2 is more complex than Test 1, since, although the stimulus sentences are de-contextualised, listeners will be more familiar with the holistic approach to categorisation employed in Test 1, which can be regarded as similar to the process of comprehension of spoken language in everyday situations (although illocutionary function can often be accessed through the use of contextual cues). This familiarity might be better described as a result of the **similar** scenario processes. A similar scenario might mean that perception of familiar intonational patterns, access to the lexis, semantic and pragmatic levels would be done without conscious control, in an automatized or routinized process, thus reducing cognitive load (*cf*. Sweller 2010).

Test 2, on the other hand, entails the use of metacognitive strategies and metalinguistic knowledge. Attentional resources are directed towards one specific part of the stimulus (the pitch change on or as from the tonic syllable), thus seemingly facilitating processing in working memory. However, another dimension of Test 2 is given by the transformational operation entailed in mapping perceived pitch changes onto a pseudo-spatial scale along an up-down cline, which in turn relies on working memory and a capacity to track how pitch changes in time and then temporally integrate such changes. This suggests that Test 2, although focused on a smaller part of the stimulus, could be considered as having a greater degree of complexity than Test 1 and demanding a heavier cognitive load.

If these dimensions are appropriate markers of task complexity, then it follows that a higher degree of complexity promotes a more effective categorisation of **new** H* L- H% patterns as falling-rising integrations of pitch, but a less accurate categorisation of **similar** H* L- L% patterns as falling interpolations. In other words, more complex cognitive activities would have different effects on performance depending on whether the scenario is **new** or **similar** within the L2LP framework. In terms of the ability to perceive the sequence H* L- H%, the higher scores obtained by the subjects in Task 2 would align with Robinson's Cognition Hypothesis, in the sense that a more complex task yields more accuracy, whereas the ability to perceive the sequence H* L- L% accurately is negatively affected by a more complex task (in the sense of activity), analogous to Skehan's Trade-off Hypothesis (2014).

It seems that the strategy demanded by an analytic test, which is not often used as a sole resource in communicative interactions, introduces a type of cognitive processing which somehow interferes with the correct categorisation of an otherwise familiar stimulus. It is likely that the subjects' interlanguage prosodic systems were undergoing a re-structuring process triggered by the introduction of a new L2 form and form-function relationship, which may have caused instability and, perhaps, U-shaped behaviour in previously more stable cognitive structures (*cf*. McLaughlin 1987).

Another striking result is the relatively high rate of errors in the categorisation of H* L- H% as 'statement' runs counter to what could be expected in view of the L2LP model. The nuclear pitch accent H* L- H% does not occur in River Plate Spanish questions. It is therefore a new form-function relationship for which a new category must be established by the non-native listener. The form itself is also a new L2 element. It is probable that those subjects who categorised H* L- H% as statement or as fall failed to perceive the H% boundary tone and interpreted it as an L% instead. One reason for this might be that the listeners resorted to the H* L- L% L1 category which acted as a filter and prevented the detection of H%.

The type of test employed may also have activated this filter. In Test 1 there was no specific perceptual attention used. If the subjects were not consciously paying attention to whether there was a rising terminal at the end of the utterance (the edge tones L-H%), perhaps this caused the assimilation of the new L2 pattern to be assimilated into the L1 H* L- L% category. This situation fits the **new** scenario in the L2LP model, where the learner perceives fewer categories than the ones existing in the L2.

A further factor favouring L1 influence may be the errors induced by the nature of instruction received (Ellis, 1994:60). O'Connor and Arnold's (1973) descriptions of each of the forms of intonation, still widely used in L2 English intonation teaching, are based on extreme pitch obtrusions, and wider-than-normal pitch ranges. So are most of the examples provided by books aimed at L2 learners (*cf*. Wells 2006). Teachers and lecturers also tend to widen the range between excursion points, and rightly so, since most students tend not to perceive certain L2 pitch differences. However, the use of exaggeration in teaching contexts, where a wider-than-normal pitch ranges tend to be used by the teacher and in teaching materials, and the absence of a training stage including exposure to and production of more challenging, authentic or semi-authentic materials, with narrower and faster pitch ranges, can induce learners to expect extreme pitch excursions which may not be available in authentic materials, where terminal rises L-H% may have short durations and may be perceived as L-L%, especially in words with one short post tonic syllable, e.g. *ice box, fit it.* Failure to detect such short-lived rising terminals may lead to a L- L% psychoacoustic categorisation.

Another dimension along which perception and categorisation of H* L- L% and H* L- H% differ might be related to the effect of alignment of L- with respect to H%, that is to say, the duration of the voicing period between the low phrase tone and the high boundary tone in minimal pairs. This interpolation is sometimes known as 'terminal rise' in the contours approach. It is well known that some English vowels have longer intrinsic duration than others. Lax vowels tend to be perceived as shorter than tense ones. Since the percept 'intonation' results from the **temporal integration**

of the relevant tones (manifested as F0 points), it is likely that some of the shorter L-H% durations went unnoticed by the listeners. House (1990) explains that a fundamental frequency variation that takes place during a given time interval will be perceived as a pitch movement only if the rate of change exceeds some minimal amount, known as 'glissando differential threshold'. The shorter the stimulus, the larger the frequency change required for a noticeable difference. Frequency variations below this threshold are perceived without pitch change (*i.e.*, they are perceived as static tones). This finding comes from research into the perception of pure tones in synthetic vowels. House (1990) refers to the lack of research of this phenomenon in speech, as certain variables that should be controlled for cannot be so in natural speech. Whether the glissando threshold is universal or language specific is not yet known, but given that a terminal rise L-H% is part of a new L2 form, it is possible that at least some listeners will have difficulty in detecting it in at least some short-lived syllable nuclei. The fact that some short terminal rises were correctly perceived and some longer terminal rises were ignored points to the competition between different prosodic cues in the categorisation, as well as the complexity of the process under study.

Figures 3 and **4** below show some of the results organised in minimal pairs. For each nuclear tonal sequence, results at Test 1 and 2 are given. Although in some cases, the longer durations in **figure 2** were categorised successfully, this was not always the case. For example, in the pairs 'bit-bit', 'fought-ford', it was the shorter duration that obtained more scores for the falling-rising category. However, the longer words, such as 'sea' and 'four', seemed to have acted as more efficient carriers of the H* L- H% configuration, and the less successful categorisation did take place on 'knot', the token which proved most difficult to identify as a falling-rising contour for all participants (notice that 'nod', with a longer duration, seems to have enabled a better categorisation, especially as falling-rising).

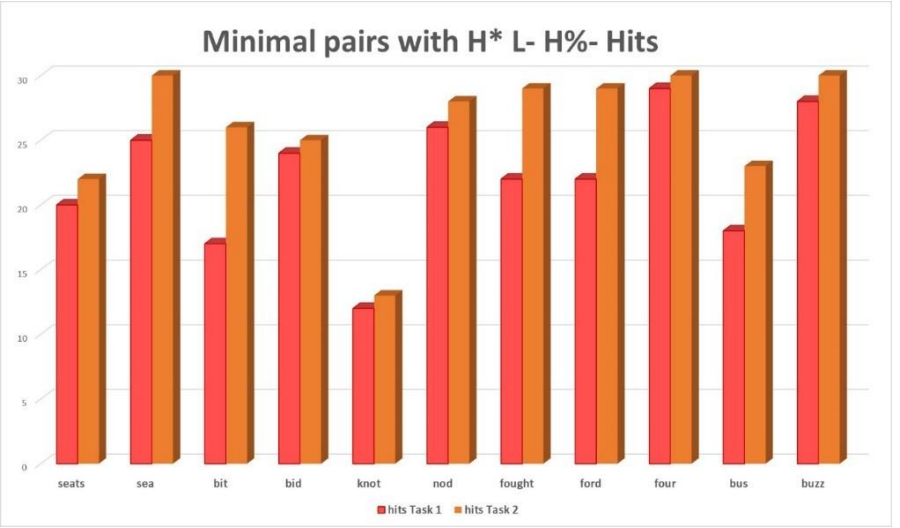

Figure 3. Correct categorisations of H*L-H% in minimal pairs in both tests.

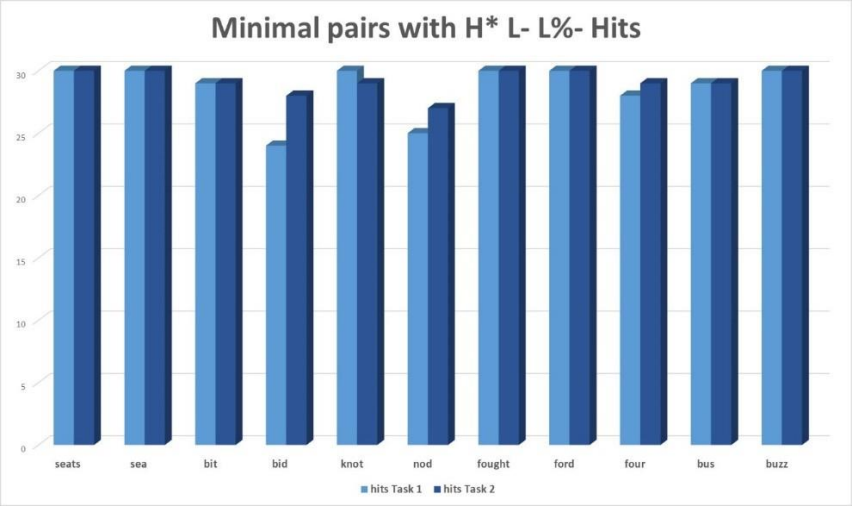

Figure 4. Correct categorisations of H*L-L% in minimal pairs in both tests.

In **figure 4**, the H* L- L% tonal configuration is shown, also arranged according to hits in minimal pairs. Besides as less variable performance

across the pairs, it can be seen that 'nod' fared worse than its longer counterpart, and 'bid,' which is longer than 'bit,' also fared worse, and not better, as it would have been expected if duration correlated directly with perception. The results are therefore not conclusive, and more research should be done in order to assess the relevance of the linguistic items used to teach this new intonational pattern.

Conclusion and suggestions for further research

The type of listening mode seems to affect the perceptual categorisation of nuclear pitch accents and the interpretation of their meanings. A listening mode with a focus on intonational forms can promote correct categorisation of H* L- H% as 'falling-rising,' but also interfere with the correct categorisation of H* L- L% as 'falling,' whereas the latter's categorisations as 'statement' is favoured by the focus on meaning. In terms of the L2LP model, listening to L2 forms in a **similar** scenario may hinder the categorisation of intonational function, but this listening mode may be a useful resource for the (more) accurate perception and categorisation of an L2 contrast at early stages (**new** scenario).

Results suggest that psychoacoustic criteria can be a useful tool in learning L2 intonation at early stages. Tonemic dictation tasks, where the intonation of speech is transcribed using tonemic marks representing the intonation patterns perceived, exploit the ability to listen to forms, *i.e.* ascending and descending pitch intervals, by means of an analytic listening strategy, and in a bottom-up process, enables the listener to arrive at the intended nuclear pitch accents. This process, in turn, leads to a meaning, which may have one or more functions, depending on the context of discourse.

If correct categorisation in a **new** scenario is promoted by analytic listening, this type of focus-on-forms activity can be beneficial in contexts of explicit instruction, when introducing new L2 intonation forms to learners. This bottom-up strategy may seem out of tune with current trends in the teaching of other L2 levels of analysis, but it should be borne in mind that a focus on forms does not preclude other teaching activities, and that not

all levels of language are necessarily learnt in the same way. What remains to be investigated is how learners can move from focus on forms on to focus on meaning and vice versa, both in terms of tests and strategies, as well as cognitive processes involved.

It is clear that the type of listening mode is only one of many factors that impinge on perception and categorisation. It seems that the number of syllables, especially post-tonic, the presence of a pretonic segment, and the duration of terminal rises (which in turn is determined mainly by the segmental composition of the syllable carrying the edge tones/terminal rise) also seem to play a role. Further research could shed light on the extent to which certain phonetic features may promote or impede accurate perception and categorisation so that ear-training materials could be carefully graded to help listeners gradually attune to small and/or short duration differentials of pitch. Finally, a related issue deserving research is whether perception of pitch differentials and their temporal integration have language-specific thresholds (as in the case of the rising terminal in H* L- H%), and to what extent perceptual training can refine correct selection and weighting of prosodic cues.

References

Arvaniti, A. (2019). Crosslinguistic variation, phonetic variability, and the formation of categories in intonation. In S. Calhoun, P. Escudero, M. Tabain & Paul W. (Eds.). *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 1-6). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Ayers, G. M. & Beckman, M. E. (1997). Guidelines for ToBI labelling (version 3.0). The Ohio State University Research Foundation. Accessed on July 6, 2020, from http://www.ling.ohio-state.edu/tobi.html

Best, C. (1995). A direct realist view of cross-language speech perception. In: W. Strange (Ed.). *Speech perception and linguistic experience* (pp. 167-200). Timonium, MD: York Press.

Boersma, P. (2003). Stochastic optimality theory. *Meeting of the Linguistic Society of America in Atlanta*, 3, p. 2003.

Boersma, P. y Weenink, D. (2019). *Praat: Doing Phonetics by Computer* (Version 6.0.47) [Computer software]. http://www.praat.org/

Braun, B., Kochanski, G., Grabe, E. & Rosner, B. S. (2006). Stable attractors in intonation: Evidence for discrete choices in English prosody. *Journal of the Acoustical Society of America*, 119 (6), pp. 4006-4015.

Brazil, D. (1997). *The Communicative Value of Intonation in English.* Cambridge: Cambridge University Press.

Council of Europe. (2001). *Common European Framework of Reference for Languages (CEFR)*. Accessed on June 30, 2020, from https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions

Cruttenden, A. (1981). Falls and rises: Meanings and universals. *Journal of Linguistics*, 17 (1), pp. 77-91.

Cruttenden, A. (2007). *Intonation*. Cambridge: Cambridge University Press.

Escudero, P. R. (2005). *Linguistic Perception and Second Language Acquisition. Explaining the attainment of optimal phonological categorization*. The Netherlands: LOT Publications.

Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In: W. Strange (Ed.). *Speech perception and linguistic experience* (pp. 233-277). Timonium, MD: York Press.

Gabriel, C., Pesková, A., Labastía, L. & Blázquez, B. (2013). La entonación en el español de Buenos Aires. En L. Colantoni & C. Rodríguez Louro(Eds.). *Perspectivas teóricas y experimentales sobre el español de la Argentina* (pp. 99-115). Madrid, Frankfurt am Maim: Iberoamericana, Vervuert.

Grabe, E., Rosner, B. J., García Albea, E. & Zhou, X. (2005). Perception of English Intonation by English, Spanish and Chinese Listeners. *Language and Speech* 46 (4), pp. 375-401.

Hadding-Koch, K. & Studdert-Kennedy, M. (1964). Intonation contours evaluated by American and Swedish test subjects. *Proceedings of the 5th International Congress of Phonetic Sciences*, pp. 326-331.

Halliday, M. A. K. (1967). *Intonation and Grammar in British English*. The Hague: Mouton.

Heeren, W., Bibyk, S., Gunlogson, C., & Tanenhaus, M. (2015). Asking or telling: real-time processing of prosodically distinguished questions and statements. *Language and Speech*, 58 (4), pp. 474-501.

House, D. (1990). *Tonal Perception in Speech*. Lund: Lund University Press.

Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent* [Doctoral dissertation]. University of Stuttgart.

Keijzer, M. (2007). *Last in first out? An investigation of the regression hypothesis in Dutch emigrants in Anglophone Canada* [Doctoral dissertation, Vrije Universiteit Amsterdam]. Research VU. https://research.vu.nl/en/publications/last-in-first-out-an-investigation-of-the-regression-hypothesis-i

Ladd, R. (1996). *Intonational Phonology.* Cambridge: Cambridge University Press.

Lindsey, G. (2017). *English after RP. Southern British Pronunciation Today*. Pilar: Tequisté.

Long, M. (1991). Focus on Form: A Design Feature in Language Teaching Methodology. In K. De Bot, R. Ginsberg & C. Kramsch (Eds.). *Foreign Language Research in Cross-Cultural Perspectives* (pp. 39-52). Amsterdam: John Benjamins.

McLaughlin, B. (1987). *Theories of second-language learning*. London: Arnold.

Mennen, I. (1998). Second language acquisition of intonation: The case of peak alignment. *Proceedings of the Linguistics Society of Chicago*, 34, pp. 327-341.

O'Connor, J. D. & Arnold, G. F. (1973). *Intonation of Colloquial English.* London: Longman.

Pierrehumbert, J. B (1980). *The phonology and phonetics of English intonation* [Doctoral dissertation, Massachusetts Institute of Technology]. DSpace@MIT. https://dspace.mit.edu/handle/1721.1/16065

Pierrehumbert, J. & Steele, S. (1989). Categories of tonal alignment in English. *Phonetica*, 46 (4), pp. 181-196.

Prince, A. & Smolensky, P. (1993). *Optimality theory: Constraint interaction in generative grammar*. Rutgers Center for Cognitive Science. http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF

Rauber, A., Rato, A., Kluge, D. & Santos, G. R. (2012). *TP: Testes / Treinamentos de Percepção* (Version 3.1) [Computer software]. http://www.worken.com.br/tp_regfree.php/

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics*, 43, pp. 1-32.

Schmid, M. (2020, May 5). Language attrition. https://l1attrition.files.wordpress.com/2019/02/english-1.pdf.

Skehan, P. (Ed.). (2014). *Processing Perspectives on Task Performance*. London: John Benjamins.

Studdert-Kennedy, M. & Hadding-Koch, K. (1964). An experimental study of some intonation contours. *Phonetica*, 11 (3-4), 175-185.

Sweller, J. (2010). Cognitive Load Theory: Recent Theoretical Advances. In J. Plass, R. Moreno & R. Brünken (Eds.). *Cognitive Load Theory* (pp. 29-47). Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511844744.004

Tench, P. (1996). *The Intonation Systems of English*. London: Cassell.

Wells, J. C. (2006). *English intonation. An introduction*. Cambridge: Cambridge University Press.

Zárate-Sández, G. (2018). "Was that a question?" Perception of utterance-final intonation among L2 learners of Spanish. In J. Levis (Ed.). *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching Conference* (pp. 76-86). Ames, IA: Iowa State University.

Appendix

List of carrier sentences. Each was presented with H* L- H% and with H* L- H%. The stimuli were randomised.

A glass bead.
Just a bit.
He made a bid.
She's going by bus.
They heard the buzz
They tied the knot.
She gave a nod.
The contest will be hard-fought.
Ask Mr. Ford.
She gave him four.
They wanted two seats.
Near the sea.
They'll fit it.
They were connected by a fitting.
She'll leave it.
She's leaving.
They brought an icebox.