THE ROLE OF INTONATIONAL CUES IN THE PERCEPTION OF DECLARATIVES AND ABSOLUTE INTERROGATIVES IN CASTILIAN SPANISH

TIMOTHY L. FACE
University of Minnesota
facex002@umn.edu

ABSTRACT

This study investigates the impact on speech perception of the four primary differences that have been observed in the intonation patterns of declaratives and absolute interrogatives in Castilian Spanish. By means of a gating experiment, stimuli of partial or complete utterances in which the F0 pattern was manipulated to contain all possible combinations of cues to sentence type were presented to listeners. For the partial utterances, listeners responded following each gate as to whether they perceived a declarative or an interrogative utterance. For complete utterances, listeners likewise responded following each gate as to the type of sentence perceived, but also rated the naturalness of the utterance. Results show that listeners achieve a high degree of accuracy in perceiving sentence type from the first intonational difference presented, that all cues to sentence type affect perception to at least some degree, that the medial F0 pattern is the strongest non-final intonational cue to sentence type, and that in spite of the strength of perception based on earlier cues, the final F0 movement has the strongest effect on perception, overriding all previous cues.

Key words: *intonation, speech perception, declaratives, absolute interrogatives*.

RESUMEN

El presente estudio pretende investigar el efecto de las diferencias entonativas entre los enunciados declarativos y los enunciados interrogativos absolutos en la percepción. En un experimento de percepción se presentaron a los participantes oraciones parciales y completas sintetizadas que contenían todas las combinaciones de pistas posibles según el tipo de enunciado (i.e. declarativo o interrogativo). Los participantes escucharon cada estímulo por partes y, después de cada parte, respondieron para identificar la oración como declarativa o interrogativa. En el caso de los estímulos que contenían oraciones completas, al escuchar la oración entera también evaluaron la naturalidad de la oración. Los resultados indican que los participantes reconocen el tipo de enunciado desde la presentación de la primera diferencia entonativa, que todas las diferencias entonativas afectan a la percepción, que el patrón entonativo en mitad de la oración tiene más influencia que otros aspectos no finales de la entonación, y que, aunque los participantes estén seguros del tipo de oración antes de escuchar su última parte, el movimiento final es tan fuerte que puede cambiar su interpretación, sin tener en cuenta las pistas anteriores.

Palabras clave: *entonación, percepción del habla, declarativas, interrogativas absolutas.*

1. INTRODUCTION

Declaratives and absolute interrogatives in Spanish are often lexically and grammatically ambiguous, with intonation and context being the factors at the listener's disposal to aid in distinguishing these two sentence types. This point can be illustrated by pairs of sentences such as those in (1). The declarative sentence in (1a) and the interrogative in (1b) employ the same words in the same order. While morphology and syntax distinguish the sentence types in the English glosses, no such distinction is present in the Spanish sentences[1].

(1) a. Compró pan en el mercado. 'He bought bread at the market.'
b. ¿Compró pan en el mercado? 'Did he buy bread at the market?'

Traditional studies of Castilian Spanish intonation have considered the final fundamental frequency (F0) movement of the utterance to be the primary intonational difference between these sentence types, with the F0 falling at the end of declaratives and rising at the end of absolute interrogatives. While some studies have also reported that the initial F0 peak of the utterance is higher in interrogatives than in declaratives, going back at least as far Navarro Tomás (1944), this has typically been seen as secondary to the final F0 pattern in differentiating declaratives from absolute interrogatives. Quilis (1993), for example, states that *El enunciado declarativo se caracteriza por la terminación descendente* 'The declarative utterance is characterized by its final fall' (428) and also that *El enunciado interrogativo absoluto…[s]e caracteriza por un final ascendente del fundamental* 'The absolute interrogative utterance is characterized by a final rise in F0' (429), distinguishing these two types of utterances intonationally only by the F0 at the end of the utterance. Kullová (1987:20) provides an example of a scholar who recognizes the higher initial F0 peak of absolute interrogatives as compared to declaratives, yet dismisses it as unimportant to communicating the difference between these two sentence types: *Al someter a un análisis musical varias oraciones españolas, llegamos a la conclusión de que no puede haber diferencias fundamentales entre la inflexión inicial…de una oración interrogativa y la de una oración afirmativa* 'After submitting several Spanish utterances to tonal analysis, we come to the conclusion that there cannot

[1] It must be noted that English also has declarative and absolute interrogative sentences which are not distinguished by word order, as in 'He bought bread at the market' and 'He bought bread at the market?' These pairs differ from their Spanish counterparts, however, in that in Spanish the interrogative is a simple request for information, a reading not possible for the English interrogative, which expresses surprise.

be fundamental differences between the initial inflexion of an interrogative utterance and that of a declarative utterance.' Yet in spite of claims of this type, there has been little investigation of the role of cues to sentence type in Castilian Spanish other than that of the final F0-movement of the utterance. Sosa (1999:216), in direct response to the quote by Kullová given above, states that *Es obvio que conclusiones de este tipo necesitan mejor fundamentación empírica basada en datos de primera mano para ser fidedignas, no partiendo de presuposiciones* 'It's obvious that conclusions of this type need better empirical grounding based on first hand data in order to be reliable, not being based on presuppositions.' Yet little empirical work to date has investigated the role of different intonational cues to sentence type in Castilian Spanish.

While traditionally the final F0 movement is considered to distinguish Castilian Spanish declaratives and absolute interrogatives, studies have shown that the intonation patterns of these utterances differ in more ways than traditionally claimed, including beyond the difference in initial F0 peak height. Face (2004), for example, reports that the intonation patterns of Castilian Spanish declaratives and absolute interrogatives show differences in F0 peak height, the presence *vs.* absence of pitch accents, the type of pitch accents used, and the direction of the utterance-final F0 movement. The existence of multiple differences in the F0 patterns between these sentence types highlights further the need for empirical investigation of the role of these intonational differences called for by Sosa (1999). Studies conducted by Sensui (1995, 2003) and Face (2005) have begun to investigate the effects of such differences on the perception of sentence type, yet none of these studies fully investigates the roles of each of the intonational differences now known to exist between Castilian Spanish declaratives and absolute interrogatives, leaving our knowledge of their roles incomplete. Nonetheless, there are indications in these studies that listeners are able to make use of intonational differences occurring earlier than the final F0 movement of the utterance in order to determine with a high degree of accuracy the type of sentence that they are hearing. The present study seeks to expand on these previous studies by investigating which of these earlier cues lead to accurate perception of sentence type.

In addition to considering which earlier cues lead to accurate perception of sentence type, the present study also considers the relative perceptual influence of each of the intonational differences between Castilian Spanish declaratives and absolute interrogatives. Even if an earlier cue is shown to lead to accurate perception of sentence type, could it be that the final F0 movement is actually a stronger cue, but that this cannot be determined in natural speech stimuli since all

cues agree with each other as to sentence type? This is, in fact, one of the limitations of the study conducted by Face (2005), and there is evidence from Dutch that the final F0 movement may override earlier cues to sentence type (van Heuven and Haan 2002). Therefore it is crucial that the experimental design of the present study allow for conflicting cues in order to address the issue of relative cue strength in perception.

The remainder of this paper is organized as follows: Section 2 presents the intonation patterns of Castilian Spanish declaratives and absolute interrogatives, focusing on the differences between the two. Section 3 contains a discussion of previous perception studies of Castilian Spanish sentence type, examining especially Face's (2005) results and how they motivate the present study. In Section 4, the experimental methods of the perception experiment carried out for the present study are presented. Section 5 presents the results of the perception experiment, and these results are then discussed in Section 6. Finally, Section 7 presents a brief summary of the paper and its conclusions.

2. INTONATION OF CASTILIAN SPANISH DECLARATIVES AND ABSOLUTE INTERROGATIVES

2.1. Declaratives

The intonation of Castilian Spanish declaratives has been very well studied (e.g. Navarro Tomás 1944, Quilis 1993, de la Mota 1995, Sosa 1999, Nibert 2000, Beckman, Díaz-Campos, McGory and Morgan 2002, Face 2002a, Hualde 2002, Ramírez Verdugo 2005, among many others), and has several characteristics that can be considered typical[2]. First, each stressed syllable is generally accompanied by a rise in F0. The rise consists of an F0 valley occurring near the onset of the stressed syllable, meaning that the F0 rise occurs during the stressed syllable. Where the F0 rise reaches its peak is dependent on the location of the F0 rise within the utterance. In pre-nuclear (i.e. non-final[3]) position, the F0 rise continues

[2] Throughout this paper the term *declarative* refers to broad focus declaratives, where no portion of the utterance is highlighted more than others (Ladd 1980).

[3] The term *nuclear* has been used in different ways by different scholars of intonation. In some cases it refers to the most prominent pitch accent, and in others to the last pitch accent.

into the syllable following the stressed syllable, reaching its peak there[4]. In nuclear (i.e. final) position, the F0 rise is reached within the stressed syllable. Finally, Spanish broad focus declaratives are characterized by downstepping, with each subsequent F0 peak achieving a lower Hertz (Hz) value than the preceding peak. A typical Spanish broad focus declarative intonation pattern, demonstrating these characteristics, is shown in figure 1 (adapted from Face 2004:67), where the stressed syllables are shaded.

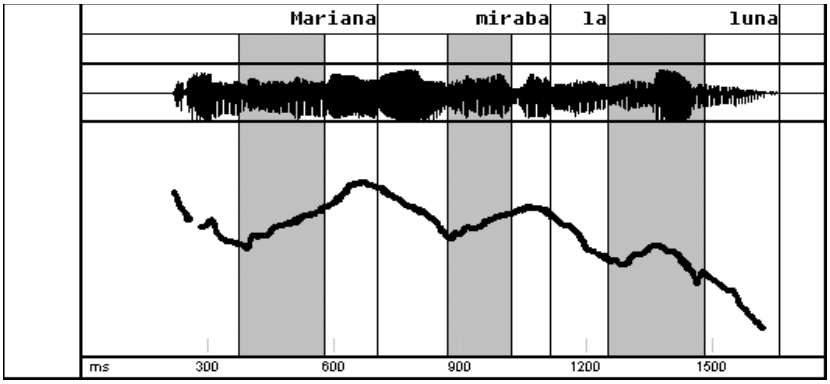

Figure 1. *Broad focus reading of the declarative* Mariana miraba la luna *'Mariana was looking at the moon'. Shading indicates the stressed syllables. Adapted from Face, 2004:67.*

2.2. Absolute interrogatives

The intonation of Castilian Spanish absolute interrogatives has received far less attention than has the intonation of declaratives. Nonetheless, there are a few

Much of this confusion is due to the fact that in some languages the most prominent pitch accent is also generally the last. This is not necessarily so in Spanish (see, for example, Face 2002a), where *nuclear* typically refers to the final accent in the phrase (with this including intermediate phrases as well as intonational phrases for those scholars that propose two levels of phrasing in Castilian Spanish).

[4] In cases of extreme tonal crowding, where two stressed syllables are adjacent to each other, the F0 rise on the first of the adjacent stressed syllables often reaches its peak within the stressed syllable in order to accommodate another F0 rise on the following syllable (Face, 2002b).

scholars that have examined the intonation of absolute interrogatives and it is possible to point out some characteristics and differences from the intonation of declaratives described above. Navarro Tomás (1944) describes the intonation of absolute interrogatives as having a high tone at the beginning of the utterance which then descends gradually throughout the utterance until it reaches the penultimate syllable of the utterance, and then rises on the final syllable of the utterance. This description indicates that words in the middle of the utterance are unaccented, showing only interpolation between the high F0 early in the utterance and the low F0 late in the utterance. Illustrations in Quilis (1993) and Ramírez Verdugo (2005) are consistent with this description[5], and Sosa (1999) and Prieto (2004) report the same, with Prieto noting that in none of her sentences did speakers produce a pitch accent other than on the initial and final words. An example of this intonation pattern is shown in figure 2 (adapted from Face 2004:70). Face (2004) reports this same intonation pattern as the most common pattern found in his study, but reports also that in some cases speakers did produce an F0 rise on the medial stressed syllable. An example of this pattern can be seen in figure 3 (adapted from Face 2004:70).

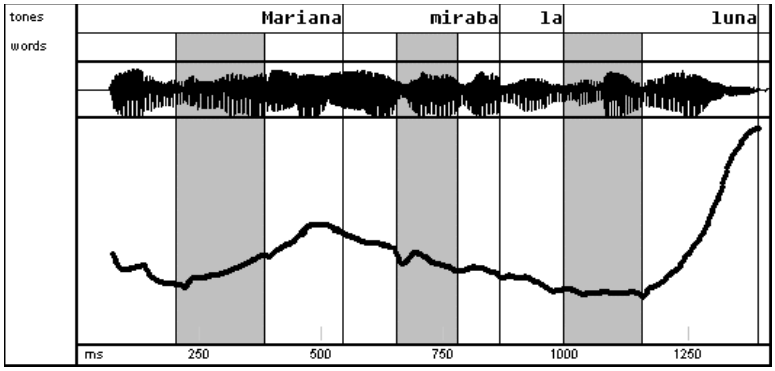

Figure 2. *Production of the absolute interrogative* ¿Mariana miraba la luna? *'Was Mariana looking at the moon?' with the medial stressed word deaccented. Shading indicates the stressed syllables. Adapted from Face 2004:70.*

[5] It must be noted that while the illustrations of absolute interrogatives in Ramírez Verdugo (2005) are consistent with this description, in her phonological analysis she does propose that there are pitch accents present in utterance-medial position.

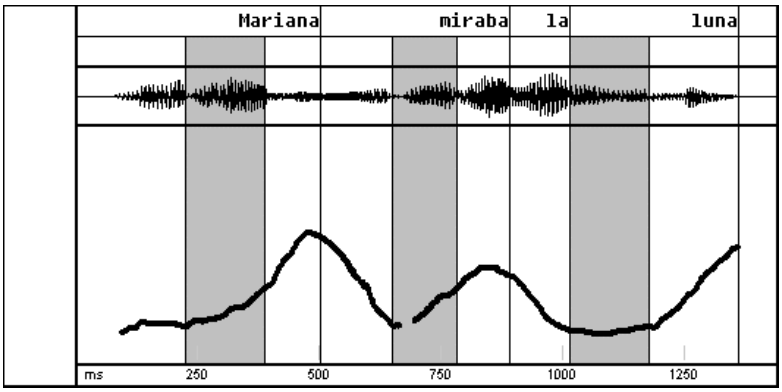

Figure 3. *Production of the absolute interrogative* ¿Mariana miraba la luna? *'Was Mariana looking at the moon?' with the medial stressed word accented. Shading indicates the stressed syllables. Adapted from Face 2004:70.*

Several scholars (e.g. Navarro Tomás 1944, Face 2004, Prieto 2004) point out that the F0 peak at the beginning of a Castilian Spanish absolute interrogative is higher than the F0 peak at the beginning of a declarative. Despite some early claims to the contrary, including by Navarro Tomás, it has been shown that the utterance-initial F0 value (as opposed to the first F0 peak) is not higher in interrogatives. Face (2004) and Prieto (2004) examined these points experimentally and found significantly higher F0 peak values for absolute interrogatives than for declaratives for all speakers in their studies, though no difference in utterance-initial F0 values was found. Face (2004) also found that the scaling difference of the initial F0 peak is the only significant difference of scaling or alignment between the initial pitch accents of declaratives and absolute interrogatives. The initial F0 rise in absolute interrogatives begins near the onset of the stressed syllable and ends in a post-tonic syllable, just as in declaratives.

While referred to briefly above in Navarro Tomás's description of the absolute interrogative intonation pattern, let us note explicitly two differences between this intonation pattern and that of declaratives that occur at the end of the utterance. First, absolute interrogatives are characterized by a low F0 throughout the final stressed syllable of the utterance, as opposed to showing an F0 rise in this syllable as do declaratives. Secondly, between the final stressed syllable and the end of the utterance the F0 rises in absolute interrogatives, while in declaratives it falls.

These differences can be seen by comparing the declarative in figure 1 with the absolute interrogatives in figures 2-3.

To summarize, the intonation pattern of Castilian Spanish absolute interrogatives differs from that of declaratives in four ways:

1. the initial F0 peak is higher in absolute interrogatives than in declaratives;

2. in medial position absolute interrogatives most commonly have no F0 rise while declaratives do;

3. during the final stressed syllable absolute interrogatives have a low F0 while declaratives have an F0 rise;

4. the F0 rises to the end of absolute interrogatives while it falls to the end of declaratives.

3. PREVIOUS RESEARCH ON CASTILIAN SPANISH SENTENCE TYPE PERCEPTION

While very few perceptual studies on any aspect of Spanish intonation have been carried out, three previous studies have examined perception of sentence type. The first two studies were small-scale studies conducted by Sensui. In the first of these studies, Sensui (1995) used filtered utterances as stimuli so that only the intonation pattern of the utterances was available to the listeners. While his results indeed demonstrated accurate perception of declarative, interrogative, and suspended utterances based only on prosody, the study does not address which differences in intonation led to this accurate perception, nor does he comment on what differences were present that could have influenced perception. In his later study, Sensui (2003) used as stimuli declarative, absolute interrogative, pronominal interrogative, and suspended utterances with either the head, body, or tail masked in white noise. His results showed that his listener was 100% accurate in determining sentence type when the tail of the utterance was available (i.e. the head or the body was masked). However, there was also a high degree of accuracy when the tail was masked, with his listener perceiving sentence type accurately 78% of the time in these cases. These results show that information earlier in the utterance did lead his listener to perceive sentence type accurately, but there are

two major limitations. First, with only one subject the results of this study cannot be considered to be a general representation of the perceptual behavior of Castilian Spanish-speaking listeners. Second, as in his previous study, there is no indication of the acoustic differences in the intonation pattern between sentence types, and therefore it is unclear what may have led to the high level of accuracy in sentence type perception.

Having determined in Face (2004) that there are four intonational differences between declaratives and absolute interrogatives in Castilian Spanish, Face (2005) set out to determine at what point in the utterance listeners correctly perceived sentence type. In other words, given the four differences in intonation between these two sentence types, which of them leads listeners to an accurate perception of whether they are hearing a declarative or an absolute interrogative? The stimuli for the experiment were productions of lexically and grammatically identical declaratives and absolute interrogatives produced by a native speaker of Castilian Spanish and divided into nine gates. For each utterance the listeners were first presented with only the first gate, then the first two gates, the first three gates, and so on until finally they heard all nine gates (i.e. the entire utterance). Four of the nine gates contained the four portions of the utterance where intonational differences between declaratives and absolute interrogatives occur, while the other five gates showed no notable differences in intonation pattern between sentence types. For each utterance, listeners were required to respond after each gate as to whether they believed the utterance to be a declarative or an interrogative. Thus listeners would respond nine times for each utterance, each time having more information upon which to base their response. The results showed that the 22 listeners were near 50% accuracy at each of the first two gates, where there was no clear difference in intonation pattern between the sentence types. However, at Gate 3, where the first F0 peak of the utterance was presented, the listeners achieved nearly 95% accuracy, indicating that the height of the first F0 peak, which is higher in absolute interrogatives than in declaratives, is an important cue to sentence type. Another significant increase in accuracy was found at Gate 5, where the declaratives had an F0 rise and the absolute interrogatives did not, with only two total errors in 660 responses. And at Gate 8, where the F0 rise during the final stressed syllable of declaratives and the low F0 through the final stressed syllable of absolute interrogatives was presented, all listeners reached 100% accuracy.

The results of Face's (2005) study show that listeners are 100% accurate in their perception of whether an utterance is a declarative or an absolute interrogative before being presented with the final F0 movement of the utterance. This is of considerable interest given the heavy importance that has been placed on the final

F0 fall of declaratives and the final F0 rise of absolute interrogatives as the primary intonational difference between these two sentence types. As a result Face (2005:62) concludes that *the utterance-final F0 movements… that are so typically associated with these sentence types carry (next to) no burden for communicating sentence type to the listener.* In addition to this apparent lack of communicative burden of the final F0 movement of the utterance, an important finding is that listeners are nearly 95% accurate in sentence type perception upon being presented with the first F0 peak of the utterance. This seems to provide an answer to the question left from Sensui's (2003) study as to what aspect of the intonation pattern early in the utterance leads to accurate sentence type perception in Castilian Spanish. Face's results indicate that the difference in F0 peak height between declaratives and absolute interrogatives is a primary cue to sentence type.

While, on the one hand, Face's (2005) results seem quite clear, it must be noted that the design of the experiment may have had an important influence on the results, leaving certain questions open for further consideration. First, the design of the perception task does not allow conclusions to be drawn as to which differences in intonation are most important in communicating sentence type. While listeners may be highly accurate in their perceptions from the very first cue to sentence type, the fact that listeners make use of this cue does not necessarily mean that it is more important than other cues. The design of the task, however, results in listeners being presented with continually more information to support their perception at the first intonational cue. If later cues were more important than the first, there would be no way to determine this. While listeners are clearly able to make use of the height of the initial F0 peak of the utterance in determining sentence type, the strong results found in Face (2005) for this cue may simply be attributable to it being the first cue presented to listeners and may say nothing about the strength of the cue other than that it is strong enough for listeners to notice it. There may be one or more stronger cues later in the utterance. In fact, in spite of the conclusion that the final F0 movement of the utterance carries little to no burden in the communication of sentence type, it is possible that the previous emphasis on this as the primary intonational difference between sentence types was indeed on target. As van Heuven and Haan (2002) found for Dutch, it is possible that the utterance-final F0 movement overrides all other intonational cues to sentence type in Castilian Spanish. The experimental design employed in Face (2005), however, does not allow for this to be determined. Second, since Face's (2005) gated stimuli were created from actual productions of both declaratives and absolute interrogatives, it is possible that differences other than intonation may have existed and could have cued listeners to sentence type. While Face considers

the possibility that duration played a role, and presents evidence that it likely did not, one must still wonder if some difference other than intonation could have come into play that would have been avoided by synthetically manipulating the F0 of the same utterance to create two stimuli: one with the F0 patterns of a declarative and one with the F0 patterns of an absolute interrogative.

4. EXPERIMENTAL METHODS

4.1. Stimuli

The present study set out to address the two concerns raised above with the study by Face (2005): it was designed to address the strength of the different intonational cues to sentence type in Castilian Spanish and to do so in a way so as to eliminate the possibility that information other than intonation could cue listeners to sentence type. In order to investigate the strength of the different intonational cues to the declarative *vs.* absolute interrogative sentence type distinction, a perception study was carried out. One of the sentences from Face (2004, 2005) that is lexically and grammatically ambiguous as to whether it is a declarative or absolute interrogative (i.e. *El marinero examina la nave* 'The sailor is examining the ship') was used in this study. This sentence had been produced as both a declarative and an absolute interrogative by speakers in Face (2004), and the production of the declarative by one of the female speakers in that study was used as the base from which stimuli for the present study were created. The pitch contour of the utterance was stylized using Praat (Boersma and Weenink 2006). The stylized contour produced pitch points able to be manipulated, and at the four places in the utterance where there is an intonational difference between declaratives and absolute interrogatives the non-essential pitch points were eliminated in order to facilitate manipulation and resynthesis of the F0 contour. Pitch points at turning points (i.e. those locations where the pitch slope changed, such as at the valley where F0 ceased to fall and began to rise) were maintained. For example, the first F0 rise of the utterance contained many pitch points, but only the pitch points at the beginning of the rise, at the peak, and at the end of the fall from the peak were maintained. Thus by manipulating the height of the pitch point at the peak, the height of the F0 rise could be manipulated and produce a natural sounding rise to the peak and fall from the peak since there would be no intervening pitch points associated with the original F0 contour (and therefore the slopes of the original rise and fall) rather than the newly created contour.

Once the F0 contour of the declarative production was stylized, it was manipulated to have the intonational characteristics of an absolute interrogative as determined by the production of the absolute interrogative version of the ambiguous sentence by the same female speaker of Castilian Spanish. For the initial F0 rise, the peak was raised to the height of the initial peak in the produced absolute interrogative. In order to remove the medial F0 rise, the pitch points creating the medial F0 rise of the declarative were removed, resulting in a steady fall in F0 between the initial F0 peak and the beginning of the last stressed syllable. To create a low F0 throughout the final stressed syllable a pitch point was added at the end of the final stressed syllable that had the same F0 value as the pitch point at the beginning of that stressed syllable, resulting in a low and flat F0 throughout the syllable. Finally, to create the rise to the end of the utterance, the final pitch point of the utterance was raised to the height of the final F0 reading of the produced interrogative.

Subsequent to the creation of a resynthesized utterance with the intonation pattern of an absolute interrogative, additional resynthesized utterances were created containing all possible combinations of the four intonational differences. That is, in addition to the utterances with completely declarative and completely interrogative intonation pattern, resynthesized utterances were created containing three characteristics of declarative intonation and one characteristic of interrogative intonation, three interrogative characteristics and one declarative characteristic, and two characteristics of each declarative and interrogative intonation, for a total of 16 resynthesized utterances.

Once the full reynthesized utterances were created, shortened utterances were created for the purpose of conducting a psycholinguistic gating experiment. Utterances would be presented to listeners in four gates, and while the full resynthesized utterances were the stimuli containing all four gates, stimuli containing one gate, two gates and three gates needed to be created. The four gates of the utterance each contained one of the four intonational differences between declaratives and absolute interrogatives. Gate 1 contained *El marinero*, and thus the initial F0 rise with peak height depending on sentence type. Gate 2 contained *examina*, which had an F0 rise in the declarative and a steady fall from the preceding F0 peak in the absolute interrogative. Gate 3 contained *la na*, (up through the stressed syllable of *nave*), which had a rise in the stressed syllable in the declarative and low and flat F0 in the stressed syllable in the absolute interrogative. For the declarative, where the F0 peak occurs in the middle of the syllable *na*, Gate 3 ended at the F0 peak so as not to include any of the final F0

fall, as this, instead of the rise, could be used in perception. Gate 4 contained *ve* (the end of the final word *nave*), which had a fall in F0 in the declarative and a rise in F0 in the interrogative. In addition, for declaratives it contained the portion of the syllable *na* following the F0 peak, meaning that it contained the entire final F0 fall. Figure 4 presents a schematic representation of the intonation pattern presented in each of the four gates when cueing a declarative utterance and when cueing an interrogative utterance. As had been done in the full resynthesized utterances (i.e. the four gate stimuli), for stimuli with one, two and three gates, each possible combination of declarative and interrogative intonation characteristics was included. This resulted in two stimuli with one gate, four stimuli with two gates, and eight stimuli with three gates. Including the 16 stimuli with all four gates, then, the total number of different stimuli was 30.

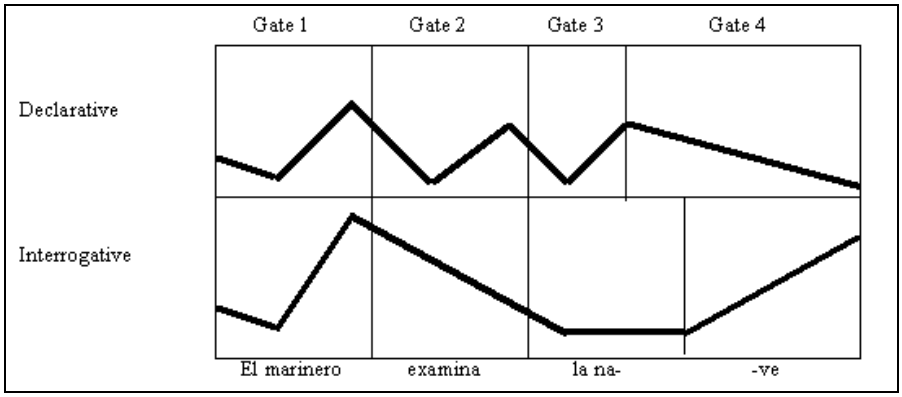

Figure 4. *Schematic representation of the intonation pattern contained within each gate in the stimuli.*

4.2. Tasks

Subjects participated in two perception tasks. For the first task they responded only to stimuli containing one, two, or three gates. These 14 stimuli were presented to subjects in random order, with each stimulus presented five times for a total of 70 trials per subject. Upon hearing each stimulus, subjects were to use the computer mouse to click on one button on the computer screen if they believed the stimulus to be from a declarative utterance and on another if they believed it to be

from an interrogative utterance. While each stimulus was presented one time automatically, the subjects had the option to click on a button at the bottom of the computer screen to hear the stimulus again, but they were limited to doing so two times for each stimulus.

For the second task, subjects responded to the full resynthesized utterances (i.e., those containing all four gates). As with the first task, these 16 stimuli were presented to subjects in random order, with each stimulus presented five times. This resulted in a total of 80 trials per subject for the second task. Upon hearing each stimulus subjects were to use the computer mouse to click on one button on the computer screen if they believed the utterance to be a declarative and another if they believed it to be an interrogative, and again they had the option of listening to each stimulus up to two additional times after the initial presentation. For this second task, after indicating whether they believed the utterance to be a declarative or interrogative, subjects were also asked to rate the naturalness of the production by clicking on one of six buttons near the bottom of the computer screen, each labeled with one of the numbers from one to six. A score of six indicated the most natural production while a score of 1 indicated the most unnatural production. The naturalness rating was included for the full utterance perception task in anticipation that subjects may well perceive declarative or interrogative based solely on whether the F0 fell or rose to the end of the utterance, regardless of the rest of the intonation contour as van Heuven and Haan (2002) found for Dutch. In this case, the naturalness rating should give an indication of the strength of the other cues even though they may not be strong enough to override the final F0 movement, with a lower naturalness rating indicating a stronger effect of cues in conflict with the cue of the final F0 movement.

4.3. Subjects

A total of 16 subjects, 10 females and 6 males, all native speakers of Castilian Spanish from central Spain, participated in this study. The subjects ranged in age from 21 to 36 years old, and each completed the perception task in a quiet office. Prior to their participation in the study, subjects were interviewed by the investigator to assure that they grew up monolingual speakers of Spanish, with any knowledge of other languages coming only through schooling rather than through living in a household or community in which another language was spoken.

5. RESULTS

5.1. Perception of sentence type: no contradictory cues

Recall that Face (2005) examined perception of sentence type in cases where all cues presented to the listeners were in agreement as to sentence type; that is, there were no contradictory cues. Let us consider these cases first, though the present study differs from that of Face (2005) in that in the present study all stimuli were created as resynthesized manipulations of the same production (a declarative utterance). The results for sentence type perception after each of the four gates are presented in table 1, with the results for intended declaratives and intended interrogatives reported separately in addition to the overall combined results.

Gate #	Perception of Intended Sentence Type					
	Declarative		Interrogative		Combined Total	
	N (out of 80)	%	N (out of 80)	%	N (out of 160)	%
1	80	100	60	75	140	87.5
2	80	100	77	96.25	157	98.13
3	78	97.5	80	100	158	98.75
4	80	100	80	100	160	100

Table 1. *Perception of the intended sentence type after each of the four gates.*

The results presented in table 1 show clearly that listeners make use of intonation from the first gate in perceiving sentence type. This confirms the finding of Face (2005), but does so while eliminating the possibility that cues other than intonation were responsible for the results. As in Face (2005), the subjects in this study make use of the height of the initial F0 peak (the only difference in intonation between intended declaratives and intended interrogatives at Gate 1) in determining sentence type. While the overall accuracy is slightly lower here than in Face (2005) – 87.5% accuracy here *vs.* 94.85% accuracy in Face (2005) – subjects in the present study are nonetheless highly accurate. It is of interest that the listeners are 100% accurate in perceiving intended declaratives, but only 75% accurate in

perceiving intended interrogatives, indicating a possible predisposition toward declaratives when there is any doubt. Even with this possible predisposition, however, listeners show remarkable accuracy in distinguishing sentence type based solely on the height of the initial F0 peak.

Once Gate 2 is presented, overall accuracy jumps to over 98%, indicating that at this point most of the doubt about intended interrogatives at Gate 1 is eliminated. Gate 2 presents the medial F0 rise in declaratives and the steady fall from the initial peak to the final stressed syllable in interrogatives (i.e. a rising pitch accent in declaratives and the lack of a pitch accent in interrogatives). As this removes most doubt about interrogatives, it seems that it may be an important perceptual cue to sentence type, though this will be considered further later on.

Upon the presentation of Gate 3, which contains an F0 rise during the final stressed syllable in declaratives and a low F0 through the final stressed syllable in interrogatives, accuracy in perceiving sentence type remains above 98%, although the two errors here are with intended declaratives instead of intended interrogatives. When the complete utterance (i.e. all four gates) is presented, listeners are 100% accurate in perceiving sentence type.

The results presented in table 1, reporting the perception of sentence type when intonational cues do not contradict each other as to which sentence type they signal, largely confirm the findings of Face (2005). The one potential difference here is that there seems to be preliminary evidence that the presence or absence of a medial pitch accent may be an important perceptual cue to sentence type. This possibility will be examined further in following subsections.

5.2. Perception of sentence type: conflicting cues

5.2.1. Perception after Gate 2

While the results presented in Section 5.1 corroborate the claims made in Face (2005) about the importance of the initial F0 peak height in the perception of sentence type, it is crucial to keep in mind that in both Face (2005) and in Section 5.1 all intonational cues to sentence type were consistent with the height of the initial F0 peak. It is only when cues come into conflict that we can determine which cues are most influential in the perception of sentence type.

Table 2 presents the results for perception of sentence type following presentation of the first two gates when the cues presented in the two gates are in conflict. That is, in stimuli where Gate 1 had a declarative intonation pattern, Gate 2 presented an interrogative intonation pattern. And in stimuli where Gate 1 had an interrogative intonation pattern, Gate 2 presented a declarative intonation pattern. Recall that at Gate 1 both sentence types have an F0 rise through the stressed syllable and into the post-tonic syllable, but the F0 peak is higher in the interrogative. At Gate 2, declaratives have a rising F0 throughout the stressed syllable and into the post-tonic syllable, while interrogatives have a steady fall between the initial F0 peak in Gate 1 and the beginning of the third stressed syllable.

Sentence Type Cue Presented		Perceived Sentence Type			
		Declarative		Interrogative	
Gate 1	Gate 2	N (out of 80)	%	N (out of 80)	%
Declarative	Interrogative	29	36.25	51	63.75
Interrogative	Declarative	51	63.75	29	36.25

Table 2. *Perception of sentence type after two gates when contradictory cues are presented in the two gates.*

Two things are immediately evident upon examining the results in table 2. First, there is less consistency in sentence type perception here than was seen for the first two gates in table 1. This is not unexpected, as the perceptions following Gate 2 in table 1 were based on two cues to the same sentence type, but in table 2 the perceptions are based on stimuli where the cues in Gate 1 and in Gate 2 were contradictory. Given conflicting information in the stimuli, one would expect less consistency in the perceptual results than in cases where there is no conflicting information.

The second thing that is evident in table 2 is that, in spite of there being less consistency here than in the perceptions in table 1, in both stimuli considered in table 2 the most common perception corresponds with the cue presented in Gate 2. This is of particular interest given how accurately subjects perceived sentence type following Gate 1. In spite of their 87.25% accuracy in perceiving sentence type

following Gate 1, when the cue in Gate 2 contradicts the cue in Gate 1, subjects perceive sentence type based on the Gate 2 cue 63.75% of the time.

The results of a 2-tailed Pearson Correlation considering all perceptions following Gate 2 (i.e. both the cases in table 2 where the cues are contradictory and those in table 1 where they are not contradictory) shows that both the Gate 1 and Gate 2 cues correlate positively with sentence type perception, and in both cases the correlations are statistically significant at the $p<0.01$ level. However, the correlation between the perceived sentence type and the Gate 2 cue (0.619) is considerably stronger than the correlation between the perceived sentence type and the Gate 1 cue (0.344). Thus in perceptions following Gate 2, both cues presented have a positive and statistically significant effect on sentence type perception, but the cue presented in Gate 2 (i.e. the medial F0 rise in declaratives *vs.* the medial steady fall between the initial F0 peak and the final stressed syllable in interrogatives) correlates more strongly with sentence type perception.

5.2.2. Perception after Gate 3

Having seen the effect of Gate 2 on listeners' perceptions of sentence type when the different cues present contradictory information, it is of interest to see whether the intonational difference presented at Gate 3 also has such an effect. Recall that at Gate 3 there is a rising F0 during the stressed syllable in declaratives, but a steady low F0 throughout the stressed syllable in interrogatives. Unlike the results for perception after Gate 2 presented in Section 5.2.1, where the sentence type cued by Gate 1 differed from that cued by Gate 2 and therefore there was one cue for each sentence type, when listeners respond following Gate 3 they do not have the simple contradiction between two cues. Rather in these cases the sentence type cued in Gate 3 will either contradict both of the cues presented in the first two gates, or it will agree with one of the first two cues but not the other. It is of interest to see whether listeners respond in favor of the sentence type cued by two of the three gates, or whether certain lone cues are strong enough to trigger sentence perception in spite of multiple cues for another sentence type.

Table 3 presents the results for perception of sentence type following presentation of the first three gates when there are contradictory cues presented to the listeners.

Sentence Type Cue Presented			Perceived Sentence Type			
			Declarative		Interrogative	
Gate 1	Gate 2	Gate 3	N (out of 80)	%	N (out of 80)	%
Declar.	Declar.	Interrog.	80	100	0	0
Declar.	Interrog.	Declar.	13	16.25	67	83.75
Declar.	Interrog.	Interrog.	24	30	56	70
Interrog.	Declar.	Declar.	50	62.5	30	37.5
Interrog.	Declar.	Interrog.	63	78.75	17	21.25
Interrog.	Interrog.	Declar.	0	0	80	100

Table 3. *Perception of sentence type after three gates when contradictory cues are presented.*

One interesting result seen in table 3 is that when Gates 1 and 2 coincide with respect to the sentence type that they cue, this is the sentence type perceived in all cases even though Gate 3 cues the other sentence type. Of course, one possible explanation for this could be that listeners are presented with two cues to one sentence type and only one cue to the other, and therefore they tend to perceive the sentence type for which there are more intonational cues. However, further data in table 3 show that this explanation cannot be maintained. In the two cases where the sentence type cued by Gates 1 and 3 coincide, but the sentence type cued by Gate 2 is in conflict with the other two cues, listeners most often perceive the sentence type cued by Gate 2, even though there are more cues for the other sentence type. This seems to add additional evidence to what has been seen above – namely that intonational cue presented in Gate 2 (i.e. the rising F0 in declaratives and steady falling F0 from the initial F0 peak in declaratives) seems to be important in the perception of sentence type. In fact, in all six combinations of cues shown in table 3, the sentence type cued by Gate 2 is always the one that is most often perceived[6].

[6] It seems strange that the sentence type cued by Gate 2 is perceived more often when both Gates 1 and 3 cue the other sentence type than when Gate 3 cues the same sentence type as Gate 2. There is no apparent explanation for this result, and it is possible that this is merely a chance result that would not repeat if the study were to be replicated.

The influence of each gate on the perception of sentence type for all perceptions following Gate 3 (i.e. both the cases in table 3 where the cues are contradictory and those in table 1 where they are not contradictory) was examined statistically by running a 2-tailed Pearson Correlation. The results, like those run for perceptions after Gate 2, show that both the Gate 1 and Gate 2 cues correlate positively with sentence type perception, and in both cases the correlations are statistically significant at the $p<0.01$ level. Also like for the results for perceptions after Gate 2, the correlation between the perceived sentence type and the Gate 2 cue (0.732) is considerably stronger than the correlation between the perceived sentence type and the Gate 1 cue (0.256), with the difference being greater here than for perceptions after Gate 2. Interestingly, the cue presented in Gate 3 has a weak negative correlation (-0.081) with the perceived sentence type. Though a fairly weak correlation, the result is nonetheless statistically significant at the $p<0.05$ level. Thus in perceptions following Gate 3, the cues presented in Gates 1 and 2 have a positive and statistically significant effect on sentence type perception, but the cue presented in Gate 2 (i.e. the medial F0 rise in declaratives *vs.* the medial steady fall between the initial F0 peak and the final stressed syllable in interrogatives) correlates much more strongly with sentence type perception. The cue presented in Gate 3 has a fairly weak negative correlation with sentence type perception, though this negative correlation does reach statistical significance.

5.2.3. Perception after Gate 4

By looking at the perception of sentence type after Gate 4 (i.e. after hearing the entire utterance) in cases of conflicting cues to sentence type, it can be seen whether the final F0 fall in declaratives and the final F0 rise in interrogatives – the intonational difference most often claimed to distinguish these sentence types, but claimed by Face (2005) to bear little communicative burden since listeners perceive sentence type accurately before hearing the final F0 movement – play an important role in communicating sentence type to the listener.

The results for perception of sentence type following presentation of all four gates when there are contradictory cues presented to the listeners are shown in table 4.

Sentence Type Cue Presented				Perceived Sentence Type			
				Declarative		Interrogative	
Gate 1	Gate 2	Gate 3	Gate 4	N (out of 80)	%	N (out of 80)	%
Declar.	Declar.	Declar.	Interrog.	1	1.25	79	98.75
Declar.	Declar.	Interrog.	Declar.	80	100	0	0
Declar.	Declar.	Interrog.	Interog.	0	0	80	100
Declar.	Interrog.	Declar.	Declar.	76	95	4	5
Declar.	Interrog.	Declar.	Interog.	0	0	80	100
Declar.	Interrog.	Interog.	Declar.	80	100	0	0
Declar.	Interrog.	Interog.	Interrog.	0	0	80	100
Interr.	Declar.	Declar.	Declar.	80	100	0	0
Interrog.	Declar.	Declar.	Interrog.	0	0	80	100
Interrog.	Declar.	Interog.	Declar.	80	100	0	0
Interrog.	Declar.	Interog.	Interog.	1	1.25	79	98.75
Interrog.	Interrog.	Declar.	Declar.	70	87.5	10	12.5
Interrog.	Interrog.	Declar.	Interrog.	0	0	80	100
Interrog.	Interrog.	Interog.	Declar.	70	87.5	10	12.5

Table 4. *Perception of sentence type after all four gates when contradictory cues are presented.*

It is apparent in table 4 that the perception of sentence type generally corresponds with the cue presented in Gate 4 – that is, the final F0 fall in declaratives or the final F0 rise in interrogatives – even when there are contradictory cues in other gates. In fact, never does less than 87.5% of the perception correspond with the Gate 4 cue, and in all but two cases it is 95% or higher. The overall rate of correspondence between the Gate 4 cue and sentence type perception in table 4 is 97.68%. It is worth noting that the only cases where there are perceptions that do not correspond with the Gate 4 cue are in cases where the Gate 2 cue contradicts the Gate 4 cue, and the only cases where less than 95% of perceptions correspond with the Gate 4 cue are cases where both the Gate 1 and Gate 2 cues contradict the

Gate 4 cue[7]. The influence of each gate on the perception of sentence type after presentation of Gate 4 (i.e. both the cases in table 4 where the cues are contradictory and those in table 1 where they are not contradictory) was examined statistically by running a 2-tailed Pearson Correlation. The results show a very strong correlation between the Gate 4 cue and sentence type perception (0.960) that is statistically significant at the $p<0.01$ level. None of the other gates show a significant correlation with perception of sentence type, and all of the correlations are quite weak, with the correlation between the Gate 2 cue and sentence type being the strongest at 0.041. Therefore upon presentation of the entire utterance, only the Gate 4 cue has a strong correlation with the perception of sentence type, and for the first time in the results presented here the Gate 2 cue does not correlate significantly with the perception of sentence type.

5.3. Naturalness ratings

While we have seen the effects of gates other than the final gate on the perception of sentence type, this was in cases where only portions of an utterance were presented to listeners (i.e. when the final gate, containing the final F0 movement of the utterance, was not presented). Anticipating the possibility of the effect reported in Section 5.2.3, that the final F0 movement overrides conflicting cues earlier in the utterance, subjects were asked to rate the naturalness of the sentence type perceived for each complete utterance (i.e. all four gates) on a six-point scale, as described in Section 4.2. Since the question to be addressed is how cues that conflict with the final F0 movement in the communication of sentence type affect the naturalness of the utterance when sentence type is perceived according to the final F0 movement, it was essential to eliminate from consideration those cases where perception of sentence type did not correspond with the sentence type communicated by the final F0 movement. With such cases removed from consideration, the naturalness ratings were considered separately for declaratives and for absolute interrogatives since it is possible that certain cues have a stronger effect on one sentence type than the other. The naturalness ratings for perceived declaratives are reported in table 5 and those for perceived interrogatives are reported in table 6 for each stimulus (i.e. each combination of the four gates).

[7] There are other cases in table 4 where the Gate 2 cue (and in some of those cases also the Gate 1 cue) contradicts the Gate 4 cue, yet perception corresponds 100% with the Gate 4 cue. Nonetheless, it is of interest that the Gate 2 cue seems to be the one that on occasion leads listeners to perceive a sentence type other than that cued by Gate 4.

Sentence Type Cue Presented				Mean Naturalness Rating	N	Standard Deviation
Gate 1	Gate 2	Gate 3	Gate 4			
Declar.	Declar.	Declar.	Declar.	5.8625	80	.38133
Declar.	Declar.	Interrog.	Declar.	5.6250	80	.66323
Declar.	Interrog.	Declar.	Declar.	3.5844	77	1.33131
Declar.	Interrog.	Interrog.	Declar.	3.6375	80	1.50311
Interrog.	Declar.	Declar.	Declar.	4.8875	80	1.00623
Interrog.	Declar.	Interrog.	Declar.	4.8250	80	1.06468
Interrog.	Interrog.	Declar.	Declar.	2.2143	70	.89916
Interrog.	Interrog.	Interrog.	Declar.	2.3143	70	.92537

Table 5. *Naturalness rating for each full utterance stimulus ending in a declarative cue and where the utterance was perceived as a declarative.*

Sentence Type Cue Presented				Mean Naturalness Rating	N	Standard Deviation
Gate 1	Gate 2	Gate 3	Gate 4			
Declar.	Declar.	Declar.	Interrog.	1.9241	79	1.05942
Declar.	Declar.	Interrog.	Interrog.	1.5125	80	.65591
Declar.	Interrog.	Declar.	Interrog.	1.9875	80	.86410
Declar.	Interrog.	Interrog.	Interrog.	4.0500	80	1.35852
Interrog.	Declar.	Declar.	Interrog.	2.4625	80	.84109
Interrog.	Declar.	Interrog.	Interrog.	2.6203	79	1.13571
Interrog.	Interrog.	Declar.	Interrog.	4.7125	80	1.27482
Interrog.	Interrog.	Interrog.	Interrog.	4.8750	80	1.16244

Table 6. *Naturalness rating for each full utterance stimulus ending in an interrogative cue and where the utterance was perceived as an interrogative.*

As can be seen in tables 5 and 6, and as is to be expected, for declaratives and for interrogatives the naturalness ratings are highest when there are no conflicting cues to sentence type[8]. When there are conflicting cues the naturalness ratings are lower. Of particular interest, however, is how the conflicting cues affect the naturalness ratings of the utterance. While a comparison of different naturalness ratings in tables 5 and 6 can provide some insight, tables 7 and 8 provide a more direct way of viewing the data of interest.

Table 7 conflates and reorganizes the data presented in table 5 for those cases where declaratives were perceived in accordance with a falling final F0 pattern presented in Gate 4. For each of the earlier gates, table 7 reports the mean naturalness ratings both for those cases where the sentence type cued by those gates agrees with the Gate 4 cue (i.e. cues a declarative) and for those cases where the sentence type cued by those gates conflicts with the Gate 4 cue (i.e. cues an interrogative while Gate 4 cues a declarative). In this way the difference between the mean naturalness ratings when each cue agrees with the Gate 4 cue can be compared to those cases where each cue conflicts with the Gate 4 cue, providing insight into which cues have the biggest influence on the naturalness of the utterance. Similarly, table 8 conflates and reorganizes the data presented in table 6 for cases where interrogatives were perceived in accordance with a rising final F0 pattern presented in Gate 4. Naturalness ratings, and the differences between them, are presented for those cases where the sentence type cued by earlier gates agrees with the Gate 4 cue (i.e. cues an interrogative) and for cases where the sentence type cued by those gates conflicts with the Gate 4 cue (i.e. cues a declarative while Gate 4 cues an interrogative.)

For each gate in both tables, the statistical significance of the mean difference in naturalness ratings is examined via an ANOVA.

[8] Comparing tables 5 and 6 it can be seen that the naturalness rating of the declaratives with all declarative cues (i.e. Declar. Declar. Declar. Declar.) is almost a full point higher than the rating of the interrogatives with all interrogative cues (i.e. Interrog. Interrog. Interrog. Interrog). It is not clear why declaratives on the whole would be ranked as more natural than the interrogatives. One possibility is a predisposition for declaratives over interrogatives (cf. Section 5.1), whatever the reason for such a predisposition may be. Another possibility is that naturalness was affected by the creation of all stimuli from what was originally a declarative utterance.

Gate	Sentence Type Cued	Mean Naturalness Rating	Difference Between Means	N	Standard Deviation	ANOVA
1	Declar.	4.6877	1.041	317	1.51164	F(1,615)=68.095, p<.001
	Interrog.	3.6467		300	1.62191	
2	Declar.	5.3000	2.3236	320	.93888	F(1,615)=605.964, p<.001
	Interrog.	2.9764		297	1.37894	
3	Declar.	4.2052	0.0471	307	1.66281	F(1,615)=0.126, p=.723
	Interrog.	4.1581		310	1.63819	

Table 7. *Naturalness rating of full utterance stimuli ending in a declarative cue by the sentence type cued at each of the first three gates, along with ANOVA results testing the significance of each gate's cue on the naturalness rating.*

Gate	Sentence Type Cued	Mean Naturalness Rating	Difference Between Means	N	Standard Deviation	ANOVA
1	Declar.	2.3699	1.3009	319	1.41683	F(1,636)=119.492, p<.001
	Interrog.	3.6708		319	1.58455	
2	Declar.	2.1289	1.7774	318	1.03521	F(1,636)=266.388, p<.001
	Interrog.	3.9063		320	1.64497	
3	Declar.	2.7743	0.4922	319	1.53138	F(1,636)=14.730, p<.001
	Interrog.	3.2665		319	1.70308	

Table 8. *Naturalness rating of full utterance stimuli ending in an interrogative cue by the sentence type cued at each of the first three gates, along with ANOVA results testing the significance of each gate's cue on the naturalness rating.*

For each of the first three gates in both of tables 7 and 8, when the cue at that gate matches the cue in the final gate (and therefore cues the same sentence type as the final gate), the naturalness rating is higher than when the cue at the earlier gate

contradicts the cue in the final gate. For declaratives in table 7, the cue at each of the first three gates has a statistically significant effect on the naturalness rating of the declarative utterance. But while all three of the non-final cues have a significant effect, it is clear through the difference between the means as well as through the F value of each ANOVA that the cue at Gate 2 (i.e. the medial F0 rise in declaratives as opposed to the steady fall of interrogatives) has the largest effect on the naturalness rating. The cue at Gate 1 (i.e. a higher F0 peak in interrogatives than in declaratives) has the second strongest effect on the naturalness rating, while the cue at Gate 3 (i.e. an F0 rise in declaratives and a low F0 in interrogatives) has the weakest effect, although the effect is still significant. The results in table 8 for interrogatives are similar in that the cue at Gate 2 has the strongest effect on the naturalness rating, followed by the cue at Gate 1. For the interrogatives, like the declaratives, the cue at Gate 3 has the smallest effect on the naturalness rating. Unlike the case of the declaratives, however, in the case of the interrogatives there is almost no difference in naturalness ratings based on the cue at Gate 3, and the difference is not statistically significant.

Summarizing the results for the naturalness ratings, we can reach a few conclusions. First, naturalness ratings are always lower when an earlier cue conflicts with the final cue than when that same earlier cue agrees with the final cue. While this is not unexpected, it is important to note as it confirms that all parts of the intonation pattern have at least some effect on perception of sentence type. Secondly, for both sentence types the naturalness ratings drop the most when the sentence type cued by the medial F0 contour (i.e. Gate 2) conflicts with the sentence type cued by the final F0 movement (i.e. Gate 4). The strong effect of Gate 2 on naturalness ratings is in line with the findings for the effects of Gate 2 on subjects' perceptions in the sentence type perception task. Third, while having less of an effect than the medial F0 contour, the height of the initial F0 peak (i.e. Gate 1) also has a significant effect on the naturalness ratings of both sentence types when the sentence type it cues conflicts with that cued by the final F0 movement (i.e. Gate 4). Finally, the F0 contour during the final stressed syllable (i.e. Gate 3) has little effect on naturalness ratings. In declaratives there is a statistically significant difference between the naturalness ratings of those cases where it cues the same sentence type as the final F0 movement (i.e. Gate 4) and those cases where it conflicts with the sentence type cued by the final F0 movement. Even so, the difference in naturalness ratings is far smaller than for the other gates. And for interrogatives, the naturalness ratings are nearly identical, and the difference not statistically significant, when the sentence type cued by the F0 pattern during the final stressed syllable agrees and when it conflicts with the sentence type cued by the final F0 movement.

6. DISCUSSION

The results of the present study add to our understanding of the role that different intonational cues play in the perception of the declarative *vs*. absolute interrogative sentence type distinction in Castilian Spanish. However, they also raise additional points worth considering, with some of these specific to Castilian Spanish and others of more general interest.

6.1. Location and meaning of elevated F0 peaks

The findings for perception of utterances where there were no contradictory cues to sentence type largely confirm the findings presented in Face (2005). After presentation of the first cue to sentence type listeners were highly accurate in perceiving sentence type at 87.5%, though slightly less accurate than the nearly 95% accuracy found in Face (2005). Nonetheless these results confirm the importance of the first cue, and do so when the stimuli are identical except for the intonation pattern since all stimuli were created from one original utterance with only F0 manipulated. It can thus be confirmed in this study that listeners make use of the height of the initial F0 peak of the utterance in their perception of sentence type. This result raises the questions of where within a Castilian Spanish utterance an elevated F0 peak may be realized and whether or not the meaning of such an elevated pitch peak is consistent across the contexts in which it can occur. While this study as well as that of Face (2005) point to the perceptual salience of the elevated initial F0 peak in absolute interrogatives, elevated F0 peaks have been reported in other locations, both in absolute interrogatives and in other sentence types. Absolute interrogatives most often have no medial pitch accent, but Face (2004) reports that some speakers do produce medial rising pitch accents in some of their productions. When these medial pitch accents are present in absolute interrogatives, the F0 peaks are higher than the peaks in the same position in declaratives. Thus throughout absolute interrogatives F0 peak height appears to cue sentence type, with the elevated peak indicating that the utterance is an absolute interrogative rather than a declarative[9].

[9] It must be pointed out that while the F0 height distinction between declaratives and absolute interrogatives is evidenced by the height of the peaks, since this height difference is seen throughout the utterance it is likely that this is a global phenomenon rather than a local phenomenon repeated at each peak.

But looking beyond absolute interrogatives, there are additional cases of distinctions in Castilian Spanish based on F0 peak height. Prieto (2004) examines the height of the initial F0 peak in various types of utterances and finds that interrogatives (including both absolute interrogatives and pronominal interrogatives) and imperatives (cf. Kvavik 1988 for Cuban Spanish, Willis 2002 for Mexican Spanish) have higher initial F0 peaks than do declaratives, and that exclamatives have an initial F0 peak higher than all other sentence types she considers[10]. This leads Prieto to propose that there is at least a two-way, and perhaps a three-way, F0 peak height distinction.

In addition to the downstepping mentioned in Section 2.1 for declaratives, Prieto, Shih and Nibert (1996), examining Mexican Spanish, reported a process of final lowering which causes a reduction of the final F0 peak of the utterances, even beyond what is expected from downstepping. Face (2002a) has found that final lowering also exists for Castilian Spanish, with the average amount of lowering from the penultimate peak to the final peak of a declarative with five stressed words being greater than the average amount of lowering by downstep over the entire earlier portion of the declarative utterance (i.e. from the first to the fourth F0 peaks). While nearly all studies of Castilian Spanish declaratives report a lower F0 peak in final position, these studies are typically based on scripted speech recorded in controlled settings. Face (2003) shows that there is much more variation in the height of the final F0 peak in spontaneous speech, and that utterances with a very prominent final F0 rise, with the peak often higher than the preceding peak, are more frequent than those showing final lowering. Further examination beyond that reported in Face (2003) indicates that this higher F0 peak in final position corresponds with new or unpredictable information, while final lowering marks information that is either repeated or that is predictable based on the context.

F0 peak height has also been examined with respect to narrow focus. García-Lecumberri (1995) reports higher F0 peaks in declaratives on words in narrow focus than on words in broad focus. While finding differences in the alignment of F0 peaks as well, de la Mota (1995, 1997) also finds higher F0 peaks in contrastive focus than in narrow focus. Yet García-Lecumberri reports inconsistency across speakers with respect to F0 peak height and its relationship to focus (as did Toledo 1989 for Buenos Aires Spanish), and Face (2000) finds that there is no significant difference in F0 peak height between cases of broad and narrow focus. Face

[10] Imperatives and exclamatives are also shown to have earlier F0 peaks than declaratives and interrogatives, and this could also play a role in distinguishing these sentence types.

(2002a, 2002b) points out that there are multiple strategies for marking narrow focus in Castilian Spanish, with F0 peak height being one of them, but not necessarily accompanying the others. When the strategies are separated out and examined separately he shows that when the F0 rise has the same shape in broad focus and narrow focus cases, the F0 peak is higher in cases of narrow focus as there is nothing else in the intonation pattern to communicate the narrow focus. In a first study on narrow focus in Castilian Spanish absolute interrogatives, Face (2006) finds that F0 peak height is used to mark narrow focus in absolute interrogatives as well.

What can be seen from this discussion is that while F0 peak height can communicate that an utterance is an absolute interrogative rather than a declarative, this same aspect of the intonational pattern can communicate other meanings as well. F0 peak height may also distinguish other types of utterances and also communicates pragmatic information such as narrow focus and new or unpredictable information. While some of the meanings of an elevated F0 peak height could be grouped together loosely as emphatic (narrow focus, highlighting new information, exclamatives), it cannot be said that absolute interrogatives are more emphatic in any way than are declaratives. F0 peak height distinctions occur in different parts of the utterance and communicate different meanings, and seem to play an important role in the intonational system of Castilian Spanish. Yet a complete understanding of how F0 peak height functions in Castilian Spanish has not yet been obtained as studies have focused more on the alignment of F0 peaks. Future studies should continue to address the nature and role of F0 peak height in marking contrasts of different types.

6.2. Cross-linguistic aspects of declarative and interrogative intonation

While higher F0 peak heights are used very early on by listeners in determining sentence type, the present study also shows the strength of the final F0 movement of the utterance, which largely overrides all other cues to sentence type in Castilian Spanish, just as van Heuven and Haan (2002) found for Dutch. Interestingly, as in the present study, van Heuven and Haan also found that one of the earlier cues that listeners made use of was the height of F0 peaks (in medial position in their study). But it is the final F0 movement that has received the most attention, and not only in experimental studies. It is often cited as a primary distinguishing factor of declaratives and interrogatives cross-linguistically. This raises the issue of whether what is found in the results of the present study is specific to Castilian Spanish or whether it is part of cross-linguistic tendencies in marking declaratives and

interrogatives intonationally. I turn to one recent study in order to consider this possibility.

Gussenhoven and Chen (2000) conducted a perceptual study with subjects who were native speakers of three languages (i.e. Mandarin Chinese, Dutch, and Hungarian) that mark interrogatives differently. Subjects were told that the utterances they would hear were from a little-known language spoken on a South Pacific island, though in reality it was made up by the experimenters. Utterances were manipulated synthetically to vary F0 peak height, F0 peak alignment, and the final F0 value of the utterance. Each of these utterances was paired with an utterance with a segmentally identical anchor stimulus that did not vary for the three factors that were varied in the other member of the pairs. Subjects had the task of identifying which member of each pair sounded most like an interrogative. Gussenhoven and Chen predicted that all subjects would make use of innate non-linguistic knowledge and select the one with a higher F0 peak, later F0 peak or higher final F0 value as the interrogative. Of course, it is also expected that subjects will show some influence of their own language experience in their judgments as well. While there was some influence of the native language of the subjects, indicating that language experience plays a role in subjects' perception strategies, the overall results are striking. The results showed that higher F0 peaks, later F0 peaks and a higher final F0 value lead Mandarin Chinese, Dutch and Hungarian speakers to perceive interrogatives in an unknown language. Since each of these languages prosodically marks interrogatives differently, and since subjects could not be familiar with the language used in the perception task, Gussenhoven and Chen argue that these three cues are universal.

While the three aspects of intonation that Gussenhoven and Chen examined perceptually were based on findings of studies on the intonation of interrogatives across languages, certainly more work is needed to determine just how universal these cues are in marking interrogatives. But at the same time, their results highlight that there are at least strong cross-linguistic tendencies in intonation and the meanings it communicates that merit further investigation. With respect to Castilian Spanish, it is of interest that two of these three cross-linguistic tendencies play such a strong role in perception of sentence type. The height of the initial F0 peak leads to a very high degree of accuracy in perceiving sentence type, and the final F0 movement (leading to a low final F0 value in declaratives and a high final F0 value in absolute interrogatives) overrides all other cues to sentence type.

6.3. Language-specific aspects of declarative and interrogative intonation

While higher F0 peaks and a higher final F0 value, in comparison to declaratives, may be cross-linguistic tendencies of interrogative intonation, these account for only two of the four intonational differences between declaratives and absolute interrogatives in Castilian Spanish. The presence of a medial F0 rise in declaratives and the lack of such a rise in absolute interrogatives is the strongest non-final cue to sentence type in the present study. For all of the other three differences there is a local intonational event in both declaratives and absolute interrogatives. In the case of the difference in initial F0 peak height, there is an F0 rise through the first stressed syllable in both sentence types. In the case of the final stressed syllable, there is an F0 rise in the stressed syllable of declaratives and a low F0 through the stressed syllable of absolute interrogatives. And at the end of the utterance there is an F0 fall in declaratives and an F0 rise in absolute interrogatives. Only at the medial stressed syllable is there a distinction between the presence *vs*. absence of a local intonational event: there is a rise through the stressed syllable in declaratives, but no local F0 movement in absolute interrogatives. In absolute interrogatives the F0 merely shows interpolation between a preceding and a following intonational event, falling from the initial F0 peak to the low F0 level in the last stressed syllable. There is no local intonational event during the medial portion of the utterance. It may be, then, that the presence *vs.* absence of an F0 rise is such a strong cue to sentence type because it involves the presence *vs.* absence of a local intonational event. This type of distinction may be more salient in general than distinctions between local intonational events. Language experience may also come into play here. Recall that while Face (2004) reports that some speakers in some productions do indeed produce a medial rise in absolute interrogatives, this occurs only in the minority of cases and has never been reported in any other study. The lack of medial rise may be so strongly associated with absolute interrogatives, then, that the distinction between this and the pattern found in declaratives is more salient than other differences.

It is also of interest that the distinction between the F0 rise in the final stressed syllable of declaratives and the low F0 throughout the final stressed syllable of absolute interrogatives has such a small influence on perception of sentence type. One might expect that this distinction would be more salient than a peak height difference since the patterns of the two sentence types are so clearly distinct in this position. In other words, it might be expected that this would be a cue similar in strength to the medial portion of the utterance just discussed. Yet there is no significant correlation between the intonation pattern in the final stressed syllable and sentence type perception in the present study. Interestingly, when full sentences were presented to listeners and naturalness ratings examined, the pattern

in the final stressed syllable significantly affected naturalness ratings for interrogatives but came nowhere near significance (p=0.723) for declaratives. It would appear from this finding that the low F0 through the final tonic syllable is more characteristic of interrogatives (and therefore naturalness is significantly affected when it is not present) than is the F0 rise in the final tonic syllable of declaratives (where naturalness is not significantly affected when the rise is not present). While Castilian Spanish declaratives are often described as having an F0 rise during the final tonic syllable, various studies have mentioned deaccenting, where a tonic syllable has no pitch accent (e.g. Face 2003, Hualde 2003) and also final lowering, which reduces the final rise sometimes to the extreme extent that there is no visible F0 rise in the pitch track (e.g. Face 2002a; also Prieto, Shih and Nibert 1996 for Mexican Spanish). Therefore the lack of a rise is attested in declaratives, which makes the lack of a rise not seem unnatural to listeners. While much less work has been done on the intonation of Castilian Spanish interrogatives to this point, all studies have shown a low F0 through the final stressed syllable prior to the final F0 rise. Since a rise in F0 during the final tonic syllable of absolute interrogatives is not attested, it is likely that it does not occur and therefore listeners find such a rise uncharacteristic of absolute interrogatives, which affects the naturalness ratings.

The results for the effects of the intonation patterns during the medial and final stressed syllables on sentence type perception in Castilian Spanish do not correspond with the cross-linguistic tendencies noted in Section 6.2 for the other intonational differences between declaratives and absolute interrogatives. Yet we see in these results how language experience can affect speech perception, in one case leading to a very strong distinction and in the other case a minimal distinction between sentence types. In the case of medial stressed syllables, the strength of association between F0 rises in declaratives and their absence in interrogatives provides a language-specific basis for a strong perceptual cue to sentence type. While at first blush there may appear to be a similar situation during the last stressed syllable, the process of final lowering, which can be extreme enough to flatten a would-be F0 rise in declaratives, means that the lack of a rise in this position is not unnatural, leading to a weak distinction between declaratives and absolute interrogatives where the F0 never rises during the final stressed syllable.

6.4. Cue strength, redundancy, and natural speech

An important issue, I believe, is to bring together the results of Face (2005) and the portions of the present study with no contradictory cues in the stimuli with the

results of the remainder of the present study where stimuli contained contradictory cues to sentence type. Neither one of these approaches on its own shows the full picture of the intonational signaling of sentence type, but bringing the two together gives us a better understanding of this issue. While the present study does indeed show that the final F0 movement is the strongest cue to sentence type, and that other cues are also relevant in the process, it took a psycholinguistic experiment involving manipulated F0 contours containing mixed cues to sentence type to come to that conclusion. As was seen in the results of Face (2005) and in the first set of results considered in the present study, listeners are highly accurate in perceiving sentence type as soon as they are presented with the very first cue. If all other cues signal the same sentence type as the first cue, there is never a chance to examine the relative strength of the various cues. In this case one comes away with the impression that the height of the initial F0 peak is the primary cue in signaling sentence type, and since complete accuracy is reached before the presentation of the final cue, the impression given is that the final cue has little role in communicating sentence type. The present study shows that this simply is not the case, as those false impressions are based on experimentation that does not allow a comparison of the strength of different cues.

Putting these two types of experimental results together we can come to three conclusions about sentence type perception in Castilian Spanish. First, as the results of the present study show, the final F0 movement of the utterance is the strongest cue to sentence type in listeners' perceptions. Second, other cues are also relevant in that they lead the listener to believe they are hearing a certain type of sentence even before the final F0 movement. Following the final F0 movement in terms of cue strength are the medial rise of declaratives *vs.* steady fall of interrogatives, the height of the initial F0 peak, and finally the F0 rise in the final tonic syllable of declaratives *vs.* the low F0 through the final tonic syllable of absolute interrogatives. Even the third strongest cue, the height of the initial F0 peak, leads to a very high level of accuracy in perception of sentence type when presented. Third, while the relative strength of cues is important because it tells us which cues are the most important in the perception of sentence type, in naturally occurring (i.e. non-manipulated) speech the strongest cues may not carry much communicative burden. Since speakers are so highly accurate in perceiving sentence type from the very first cue, in natural speech where the cues do not contradict each other, but rather reinforce each other through redundancy, the two strongest cues do not carry much communicative burden. In fact, as claimed in Face (2005), the final F0 movement, shown in the present study to be the strongest cue to sentence type, carries (little to) no communicative burden since listeners are 100% accurate in perceiving sentence type before being presented with the final F0

movement. Thus the cues that are strongest may rarely actually be the ones that lead the listener to perceive sentence type, but rather confirm the perception that has already been made.

7. CONCLUSION

The present study has shown the final F0 movements of the declarative and the absolute interrogative is the strongest intonational cue to distinguishing these two potentially ambiguous sentence types in Castilian Spanish. On the one hand this confirms long-held beliefs that the final F0 movement is what distinguishes the intonation patterns of these sentence types. On the other hand, however, the results of the present study show that the situation is not so simple. By placing cues to different sentence types in the same resynthesized utterance, the present study allows for a direct comparison of the strength of the different cues. The results show that while the final F0 movement is the strongest cue and overrides earlier cues to sentence type, those earlier cues do play a role. When the final F0 movement is not presented to listeners, listeners are still able to fairly consistently agree on sentence type based on the earlier cues. And when the final F0 movement is presented to listeners, the earlier cues have a significant effect on the listeners' rating of the naturalness of the stimuli. The medial rising pitch accent in declaratives *vs.* the steady fall from the first F0 peak to the final tonic syllable of interrogatives, only recently recognized as an intonational difference between these two sentence types, was shown to be the second strongest cue to this sentence type difference. This cue is the primary cause of sentence type perception when the final F0 movement is not presented to listeners, and is the strongest factor affecting the naturalness rating of the utterance when the final F0 movement is presented. The height of the initial F0 peak is the third strongest cue, having a significant effect on naturalness ratings, while the F0 pattern of the final tonic syllable (i.e. a rising F0 in declaratives *vs.* low F0 in absolute interrogatives) is much weaker than all other cues, though it does have a significant effect on the naturalness ratings of interrogatives but not declaratives. So while the final F0 movement may be the strongest cue to sentence type, all four intonational cues to the declarative *vs.* absolute interrogative sentence type difference come into play and, with the exception of the pattern during the final target syllable in declaratives, each contributes significantly to the perceived naturalness of the utterance.

As discussed in Section 6, the results of the present study appear to highlight cross-linguistic tendencies of the intonational differences between declaratives and

interrogatives. Future study is needed, on a variety of languages, to determine to what extent cross-linguistic tendencies may be considered truly universal.

Additional research should also consider further the role of multiple intonational cues to a single contrast. The present study shows that there are four differences in the intonation patterns of Castilian Spanish declaratives and absolute interrogatives and that three of these have a significant impact on speech perception. Formal phonology often insists on recognizing distinctive *vs.* redundant aspects of a contrast. Can the multiple cues seen here be so simply divided into these categories? It appears that this would be problematic given that listeners make use of less strong cues (those that could be considered redundant) to accurately perceive sentence type before being presented with the final F0 movement that is the strongest and most categorically used cue in speech perception (and therefore could be considered distinctive).

While a formal analysis of intonation patterns is beyond the scope of the present study, the present study also raises additional questions that will be of interest to formal phonological analyses, such as those couched within the Autosegmental-Metrical (AM) theory of intonational phonology. One issue raised by the results of this study, as well as by the results of Face (2005), is how to represent the contrastive role of differing F0 peak heights. The only phonological distinction in pitch level within the AM theory is that between high tones (H) and low tones (L). When peak height alone is contrastive (i.e. there are no alignment differences in the pitch accent), as in the case of the initial pitch accent in declaratives and absolute interrogatives in Castilian Spanish, the theory has no manner of accounting for this contrast in peak height. A related issue is what type of intonational unit would account for the higher F0 peak in absolute interrogatives. Since a higher F0 peak marks interrogativity, an aspect that cannot be seen as a characteristic of an individual lexical item nor as a marker of a phrase edge, how can it be represented given that intonational phonology represents tonal events at specific locations (such as stressed syllables and phrase edges)? Moving away from issues of F0 peak height, there are other theoretical issues raised by the results of the present study. For example, how can a formal theoretical analysis represent the different intonation patterns of declaratives and absolute interrogatives and do so in a way in which each phonological unit (e.g. pitch accent, phrase accent) communicates a specific meaning? As the AM theory is indeed one of phonology (as opposed to phonetics), and is therefore interested in representing and explaining the meaningful contrasts in the intonational system, the multiple cues to the distinction between sentence types will certainly prove challenging. We have already seen that this is a challenge for F0 peak height,

which communicates different meanings in Castilian Spanish, and it is certainly the case that finding consistent meanings for other phonological units could prove just as challenging. As an illustration of this point, consider the medial intonation pattern. It was shown here that the presence or absence of a medial F0 rise is of considerable significance in the perception of sentence type. Yet the medial pitch accent in declaratives cannot be claimed to carry the meaning of *declarative*, as the same pitch accent is used in other types of utterances, including in initial position in absolute interrogatives. And the lack of a pitch accent is the absence of a phonological unit, and therefore there is no unit present in medial position in absolute interrogatives to communicate interrogativity. Yet in some way these meaningful distinctions in the intonation patterns need to be incorporated into a complete phonological analysis. As new types of data become available, such as the perceptual data contributed by the present study, scholars working on the phonological analysis of Castilian Spanish intonation, as well as on theories of intonational phonology more generally, will need to address issues such as those mentioned here. Only in this way may future phonological analyses adequately represent the complexities of the intonational system.

ACKNOWLEDGMENTS: I would like to thank two anonymous reviewers whose comments have led to improvements in this paper, and especially to a clearer presentation.

8. REFERENCES

BECKMAN, M.; M. DÍAZ-CAMPOS; J. T. MCGORY and T. A. MORGAN (2002): «Intonation across Spanish, in the Tones and Breaks Indices framework», *Probus*, 14, pp. 9-36.

BOERSMA, P. and D. WEENINK (2006): «Praat: Doing phonetics by computer (version 4.4)» [Computer program], http://www.praat.org.

DE LA MOTA, C. (1995): *La representación gramatical de la información nueva en el discurso*, doctoral dissertation, Universitat Autònoma de Barcelona.

DE LA MOTA, C. (1997): «Prosody of sentences with contrastive new information in Spanish», in A. Botinis, G. Kouroupetroglou and G. Carayiannis (eds.): *Intonation: Theory, models and applications. Proceedings of an ESCA Workshop*, pp. 75-78.

FACE, T. L. (2000): «Prosodic manifestations of focus in Spanish», *Southwest Journal of Linguistics*, 19, pp. 45-62.

FACE, T. L. (2002a): *Intonational marking of contrastive focus in Madrid Spanish*, Munich, Lincom Europa.

FACE, T. L. (2002b): «When push comes to shove: Tonal crowding in Madrid Spanish», *Linguistic Association of Korea Journal*, 10 (1), pp. 77-100.

FACE, T. L. (2003): «Intonation in Spanish declaratives: Differences between lab speech and spontaneous speech», *Catalan Journal of Linguistics*, 2, pp. 115-131.

FACE, T. L. (2004): «The intonation of absolute interrogatives in Castilian Spanish», *Southwest Journal of Linguistics*, 23 (2), pp. 65-79.

FACE, T. L. (2005): «F0 peak height and the perception of sentence type in Castilian Spanish», *Revista Internacional de Lingüística Iberoamericana*, 2 (6), pp. 49-65.

FACE, T. L. (2006): «Narrow focus intonation in Castilian Spanish absolute interrogatives», *Journal of Language and Linguistics*, 5, pp. 295-311.

GARCÍA-LECUMBERRI, M. L. (1995): *Intonational signalling of information structure in English and Spanish: A comparative study*, doctoral dissertation, University of London.

GUSSENHOVEN, C. and A. CHEN (2000): «Universal and language-specific effects in the perception of question intonation», in *Proceedings of ICSLP 2000*, pp. 91-94.

HUALDE, J. I. (2002): «Intonation in Spanish and the other Ibero-Romance languages: Overview and status quaestionis», in C. Wiltshire and J. Camps (eds.): *Romance phonology and variation*, Amsterdam, John Benjamins, pp. 101-116.

HUALDE, J. I. (2003): «El modelo métrico y Autosegmental», in P. Prieto (ed.): *Teorías de la entonación*, Barcelona, Ariel, pp. 155-184.

KULLOVÁ, J. (1987): «Algunos aspectos de los medios entonativos en español», *Revista de Filología Española*, 67, pp. 19-34.

KVAVIK, K. H. (1988): «Is there a Spanish imperative intonation?», in R. M. Hammond and M. Resnick (eds.): *Studies in Caribbean Spanish dialectology*, Washington, DC, Georgetown University Press, pp. 35-49.

LADD, D. R. (1980): *The structure of intonational meaning*, Bloomington, Indiana University Press.

NAVARRO TOMÁS, T. (1944): *Manual de entonación española*, New York, Hispanic Institute in the United States.

NIBERT, H. J. (2000): *Phonetic and phonological evidence for intermediate phrasing in Spanish intonation*, doctoral dissertation, University of Illinois at Urbana-Champaign.

PRIETO, P. (2004): «The search for phonological targets in the tonal space: H1 scaling and alignment in five sentence-types in Peninsular Spanish», in T. L. Face (ed.): *Laboratory approaches to Spanish phonology*, Berlin, Mouton de Gruyter, pp. 29-59.

PRIETO, P.; C. SHIH and H. NIBERT. (1996): «Pitch downtrend in Spanish», *Journal of Phonetics*, 24, pp. 445-473.

QUILIS, A. (1993): *Tratado de fonología y fonética españolas*, Madrid, Gredos.

RAMÍREZ VERDUGO, M. D. (2005): «Aproximación a la prosodia del habla de Madrid», *Estudios de Fonética Experimental*, 14, pp. 310-326.

SENSUI, H. (1995): «Percepción de la entonación interrogativa del español: un estudio experimental», *Sophia Lingüística*, 38, pp. 1-23.

SENSUI, H. (2003): «A pilot case study on sentence pattern perception of Spanish», in *Proceedings of the 15th ICPhS*, pp. 1731-1733.

SOSA, J. M. (1999): *La entonación del español*, Madrid, Cátedra.

TOLEDO, G. A. (1989): «Señales prosódicas del foco», *Revista Argentina de Lingüística*, 5, pp. 205-230.

VAN HEUVEN, V. J. and J. HAAN (2002): «Temporal distribution of interrogativity markers in Dutch: A perceptual study», in C. Gussenhoven and N. Warner (eds.): *Laboratory phonology 7*, Berlin, Mouton de Gruyter, pp. 61-86.

WILLIS, E. W. (2002): «Is there a Spanish imperative intonation revisited: Local considerations», *Linguistics*, 40, pp. 347-374.