JOURNAL ARTICLE

The Importance of a Distributional Approach to Categoriality in Autosegmental-Metrical Accounts of Intonation

Francesco Cangemi and Martine Grice
IfL Phonetik, University of Cologne, DE
Corresponding author: Francesco Cangemi (fcangemi@uni-koeln.de)

When annotating a speech signal using an autosegmental-metrical model of intonation, transcribers associate portions of the F_0 contour with labels from a finite inventory of tonal categories. In the models we are concerned with here, these categories have the status of phonological units (phonological form), bridging the intrinsic variability of the speech signal (substance) with the intrinsic fuzziness of post-lexical function (meaning). This, together with the relatively small size of the label inventory, precludes a one-to-one relationship between form and substance, and/or between form and function. A Neapolitan Italian corpus of read speech is used to investigate the distributional properties of two pitch accents that have been studied extensively with respect to substance (the alignment of F_0 peaks) and meaning (sentence modality). Although there is a general consensus that peaks in this variety are aligned earlier in declaratives than in interrogatives, evidence is provided of contexts in which the converse is true, i.e., in which interrogative peaks are even earlier than their declarative counterparts. In this respect, interrogatives have a richer internal structure than declaratives. We argue that differences in how variably a prosodic category is encoded can be dealt with in an intonation transcription system, as long as this system relates phonological form (the choice of pitch accent in this case) both to phonetic substance and to meaning in a transparent way.

1 Introduction

1.1 Background

Intonation transcription within the autosegmental-metrical framework entails the use of discrete and *symbolic labels*, which in most cases (e.g., Jun, 2005) refer to *phonological units*. When opting for a particular label, human transcribers rely on their interpretation of both phonetic substance and meaning.[1] As a result, labels used in intonation transcription refer to phonological units bridging the intrinsic variability of the speech signal (substance) with the intrinsic fuzziness of postlexical function (meaning). This, together with the small size of the label inventory, precludes a one-to-one relationship (mapping) between form and substance, and/or between form and function (e.g., Nolan, 2008).

But what is the use of a transcription practice that does not employ any such mapping? One solution to the problem of mapping is to report speakers' and listeners' preferences— or most typical behaviour—in terms of percentages (Baumann et al., 2006; Schafer et al., 2000, amongst others), e.g., meaning A is expressed by category X 80% of the time,
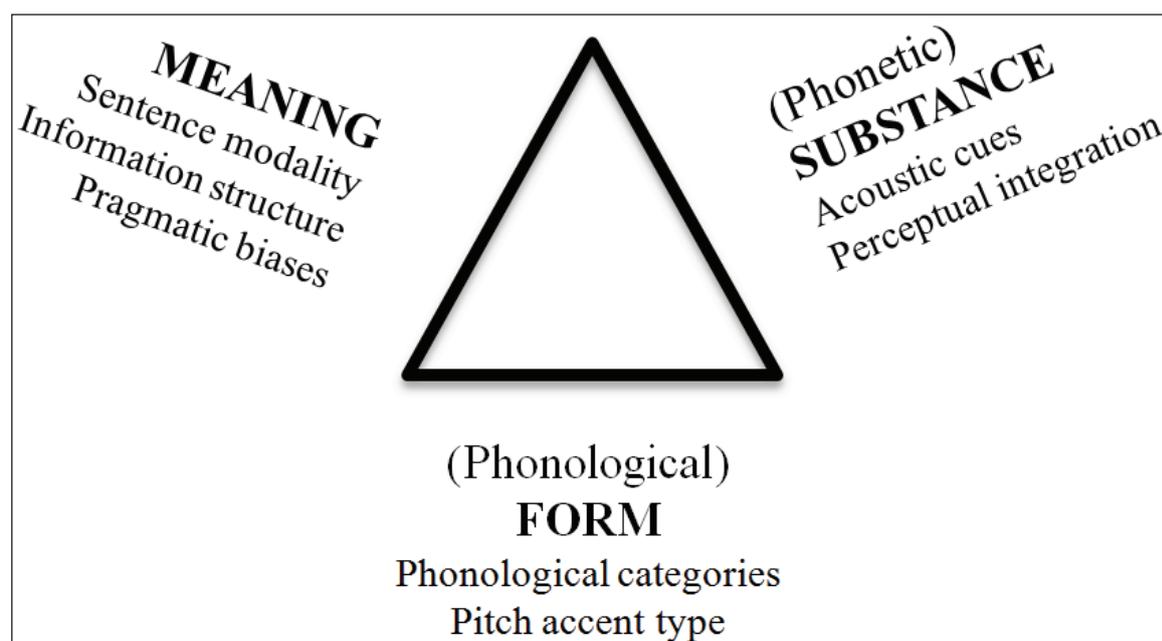
---

[1] This is true of any intonation transcription system, although priorities vary. For instance, within the British school, Crystal prioritizes substance and sees intonation as "the product of the interaction of features from different prosodic systems—tone, pitch-range, loudness, rhythmicality and tempo in particular" (Crystal, 1975, p. 283), whereas Halliday (1967) emphasizes meaning, in accordance with his understanding of intonation as a system within the grammar of English. For recent views on the importance of meaning in intonation transcription, see also Arvaniti (2016) and Cole and Shattuck-Hufnagel (2016).

Art. 9, page 2 of 20

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

this same meaning is expressed by category Y 15% of the time and by category Z 5% of the time. Such an approach employs what has recently been referred to as statistical gradience (Ladd, 2014).

Another solution—the one we are primarily concerned with in this paper—is to document the variability of phonetic parameters *within* the proposed categories (physical gradience; Ladd, 2014). We refer to this as a *distributional* approach, reflecting the importance of how the various phonetic parameters are distributed within a phonological category. If implemented with full awareness, such an approach would provide an ideal frame for the incorporation of recent developments in our understanding of the relationship between phonetics and phonology (Pierrehumbert et al., 2000) and of the dynamics between the continuous and discrete poles of linguistic knowledge (Gafos & Benus, 2006). Moreover, it would make the concept of mapping superfluous (Ohala, 1990), since the three dimensions of meaning, form, and substance would not be separable in the first place. Just as form and meaning have been traditionally linked in linguistic theory (de Saussure, 1916), a model of intonation requires a third dimension, substance, that relates equally to both form and meaning (see also Cole & Shattuck-Hufnagel, this issue). **Figure 1** sketches three dimensions as sides of the same structure (here a triangle).

1.2 Rationale

Whereas in traditional generative phonology categories are defined by the presence or absence of certain features, in a distributional approach phonological categories can be thought of as clusters in a multidimensional phonetic space (see Coleman, 2003, for a discussion in terms of contrasts, rather than categories). Such clusters can differ as to their internal structure, for example in terms of the presence or absence of sub-clusters (a sub-cluster can be seen as corresponding to an allophonic variant), or of their degree of compactness (the less compact they are, the more variation there is across individual tokens). In fact, in such an approach categories are *expected* to exhibit differences in internal structure—and such differences are in turn expected to bear on the functioning of the

Figure 1: Form, substance, and meaning in intonation transcription. Three sides of the same triangle.

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Art. 9, page 3 of 20

system as a whole. This includes how easily categories are acquired and accessed, and how prone they are to be redefined across time. In this sense, variability in encoding is seen as a resource for insights into category structure. This view is complementary to the one put forth by Cole and Shattuck-Hufnagel (this issue), in which prosodic transcription methods capitalizing on prosodic variability are proposed.

In the following, we focus on differences in internal structure across two intonational categories in Neapolitan Italian. This variety of Italian has been studied extensively (see D'Imperio, 2002; Cangemi, 2014, and references therein), especially with respect to the distinction between two pitch accents (form), the alignment of F_0 peaks (substance), and sentence modality (meaning). A corpus of read speech (Section 2.1) is used to investigate distributional properties at three levels of granularity:

First, we explore overall measures of dispersion in the fundamental frequency contours across sentence modalities (Section 2.2). We show that, independently of focus placement, Interrogatives display more variable contours than Declaratives, and that this is not an artefact of durational differences. Here we relate meaning to substance.

Second, we look at sub-clustering within each sentence modality (Section 2.3). This is done by looking at phonetic variability in the encoding of this functional contrast. There are already indications in other varieties of Italian that (polar) interrogatives have a more complex internal structure than declaratives. For instance, in Bari Italian, the bias and the expectations of the speaker when asking the question can have an effect on both the pitch accent and the boundary tone (Grice & Savino, 2003; Savino, 2014a; Savino & Grice, 2011). Moreover, whereas declaratives consistently show final falls in all varieties of Italian (Gili Fivela et al., 2015), polar interrogatives display either final rises or final falls. Differences in the final boundary tone are found not only across regional varieties (see Savino, 2012, for a recent comprehensive overview), but also within a single variety, perhaps as a function of speaking style (see Grice et al., 1997, for Bari Italian). The lack of a 1:1 mapping between final F_0 movement and sentence modality is common across a range of languages besides Italian, such as German (Kohler, 2004; Kügler, 2003) and Swedish (House, 2005). In Italian, however, intonation bears the functional load of distinguishing between the two sentence modalities, there being no morphosyntactic markers of interrogativity, such as subject-verb inversion or question particles. Thus, especially in the absence of a disambiguating context, if the final F_0 is falling, there needs to be a distinction in intonation in an earlier position for the utterance to be interpreted as a question. Examining our corpus of read speech, we find that interrogatives are indeed encoded with either final rises or final falls, indicating a difference in (internal) structural complexity between declaratives and interrogatives. Here we relate form to meaning.

Third, we focus on variability *within* intonational categories (Section 2.4). This is done by investigating peak alignment within and across pitch accents (relating form to substance). Niebuhr et al. (2011) have shown a great degree of variability across speakers in the encoding of pitch accent contrasts, referring to, for example, 'shapers' and 'aligners', where the shape of a pitch movement can be used instead of the alignment of a peak to signal category membership. Here we explore whether variability of peak alignment is equally great across the two pitch accent types we investigate, and show that peaks are aligned more variably in interrogatives than in declaratives.

In the first part of the corpus analysis (Section 2.2) we thus explore sentence modality and focus placement jointly, by measuring the variability of F_0 contours in early, medial, and late focus utterances. In all cases, interrogatives are shown to have more variable F_0 contours than declaratives. In the second part (Sections 2.3–4), we concentrate on the last pitch accent in late focus declaratives and interrogatives. In these cases the focus is on the

Art. 9, page 4 of 20

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

final word in the phrase. This more local analysis yields results which are immediately comparable to those from the exploration of global F_0 contours, with interrogatives showing richer internal structure than declaratives, due to sub-clustering and higher dispersion.
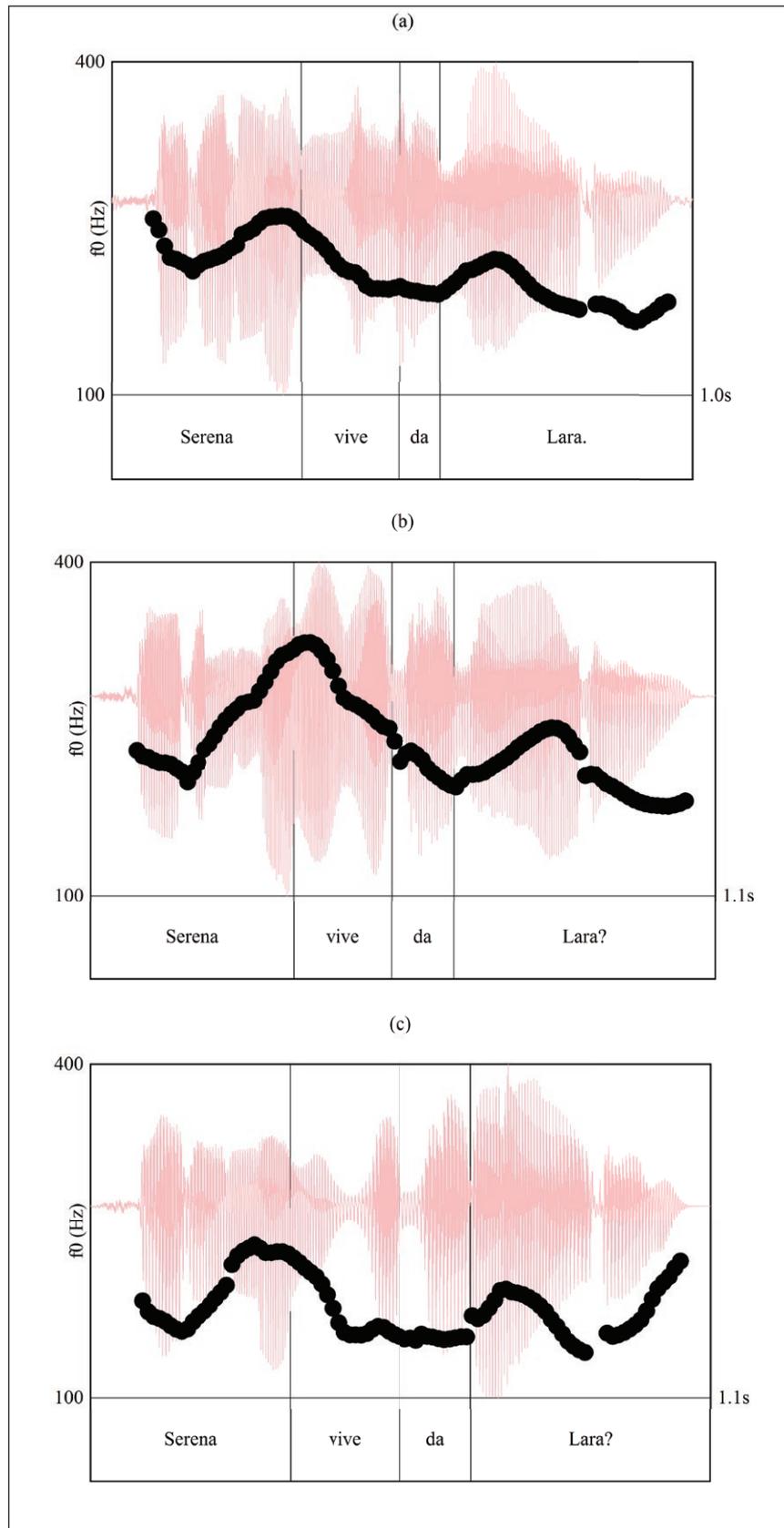
2 Corpus study

2.1 Material

Our hypotheses on differences in internal structure are tested on the *Danser* corpus (Cangemi, 2014, sec. 4.2; Cangemi & D'Imperio, 2013), which features read speech from 21 native speakers of the Neapolitan Italian variety (aged 20-25). Recordings were carried out in a sound treated booth at the University of Naples "Federico II" Interdepartmental Research Centre for Signal Analysis and Synthesis (CIRASS), using an AKG MicroMic C520 head-mounted microphone connected to a personal computer running *Audacity* (Audacity Development Team, 2006) through a Shure X2u adapter. Stimuli were prompted on a computer screen using *Perceval* (André et al., 2003).

The 21 subjects uttered 3 randomized repetitions of 6 contextually determined prosodic variants of 2 sentences after silently reading a contextualization paragraph. The sentences shared the number and structure of syllables, stress position, and syntactic structure, according to the template $[CV.C\grave{V}.CV]_s$ $[C\grave{V}.CV]V$ $[CV\#C\grave{V}.CV]_{io}$, as in *Serena vive da Lara* 'Serena lives at Lara's'. Contexts presented visually (see **Table 1** and the Appendix) induced one of the six combinations of corrective focus placement (on Subject, Verb, or Indirect Object) and sentence modality (Declarative or Polar Interrogative). For example, the context *Serena vive da Marina?* 'Does Serena live at Marina's?' was used to elicit indirect object-focussed declarative utterances of *Serena vive da Lara*.

The resulting 756 utterances were isolated from the recording sessions using *PRAAT* (Boersma & Weenink, 2008) and force-aligned at the phone level using *ASSI* (Cangemi et al., 2011). Examples are provided in **Figure 2**. Phone durations and fundamental frequency contours were extracted with *PRAAT* and analysed with *R* (Bates et al., 2014; Fox & Weisberg, 2011; R Development Core Team, 2008). Fundamental frequency contours were first time-normalized by extracting F_0 values at exactly 50 equally spaced points from each utterance, then Gaussian smoothed (bandwidth 10 Hz).

	Declarative	**Interrogative**
Subject-focus	*Tua zia ti chiede se è Ramona che vive da Lara.* *Tu rispondi:* Your aunt asks you if it's Ramona who lives at Lara's. You reply:	*Una delle tue cugine vive da Lara, ma non ricordi quale, quindi chiedi:* One of your cousins lives at Lara's, but you don't remember which one, so you ask:
Verb-focus	*Un amico ti chiede se Serena adesso lavora da Lara.* *Tu rispondi:* A friend asks you if Serena now works at Lara's. You reply:	*Serena passa molto tempo con Lara, ma non ricordi perché, quindi chiedi:* Serena spends a lot of time with Lara, but you don't remember why, so you ask:
Object-focus	*Tua sorella vuole sapere se Serena vive da Marina.* *Tu rispondi:* Your sister wonders whether Serena lives at Marina's. You reply:	*Serena è andata a vivere da un'amica, ma non ricordi chi, quindi chiedi:* Serena has moved in at a friend's, but you don't remember who, so you ask:

Table 1: Contexts for the elicitation of the six focus/modality combinations for the sentence *Serena vive da Lara* ('Serena lives at Lara's' / 'Does Serena live at Lara's?').

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Art. 9, page 5 of 20

Figure 2: Spectrogram and fundamental frequency traces with orthographic word-level segmentation for three object-focus utterances of the sentence *Serena vive da Lara*, as (a) declarative, (b) interrogative with final fall, and (c) interrogative with final rise. This audio content is available at: http://dx.doi.org/10.5334/labphon.28.wav2a, http://dx.doi.org/10.5334/labphon.28.wav2b, and http://dx.doi.org/10.5334/labphon.28.wav2c.

Art. 9, page 6 of 20

Cangemi and Grice: The Importance of a Distributional Approach to Categoriality in Autosegmental-Metrical Accounts of Intonation

2.2 Macroscopic analysis (F_0 across utterances): dispersion

A first indication that the degree of variability in realization might be different across sentence modalities comes from the mere visualization of superposed utterance-long time-normalized F_0 contours, with sentence modality presented separately. As **Figure 3** shows, when contours are pooled across focus placement conditions, interrogatives (right panel) show less homogeneous realizations across speakers, sentences, and repetitions than declaratives (left panel) do. This is confirmed by a Levene's test for homogeneity of variance ($p < 0.001$) run on contours sampled with 50 equally spaced points in time. An F-test further indicates that variance in interrogatives is 15% higher than in declaratives.

The effect holds when the three focus placement conditions are evaluated separately. Interrogatives with initial, medial, or final contrastive focus have more variable realizations than declaratives with the same focus placement. This is illustrated in **Figure 4**, and confirmed by further Levene's tests (all $p < 0.001$).
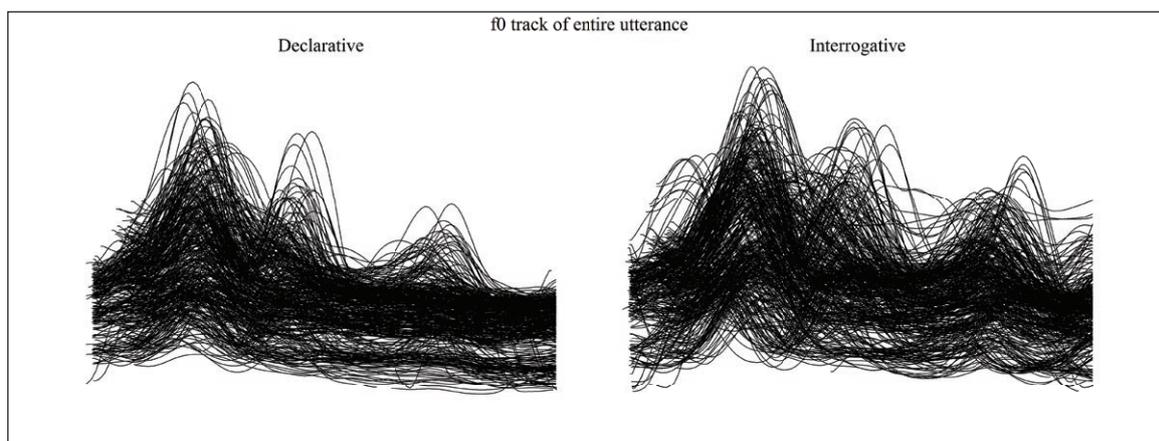
Figure 3: Utterance-long time normalised F_0 contours. Vertical range 75–450 Hz.
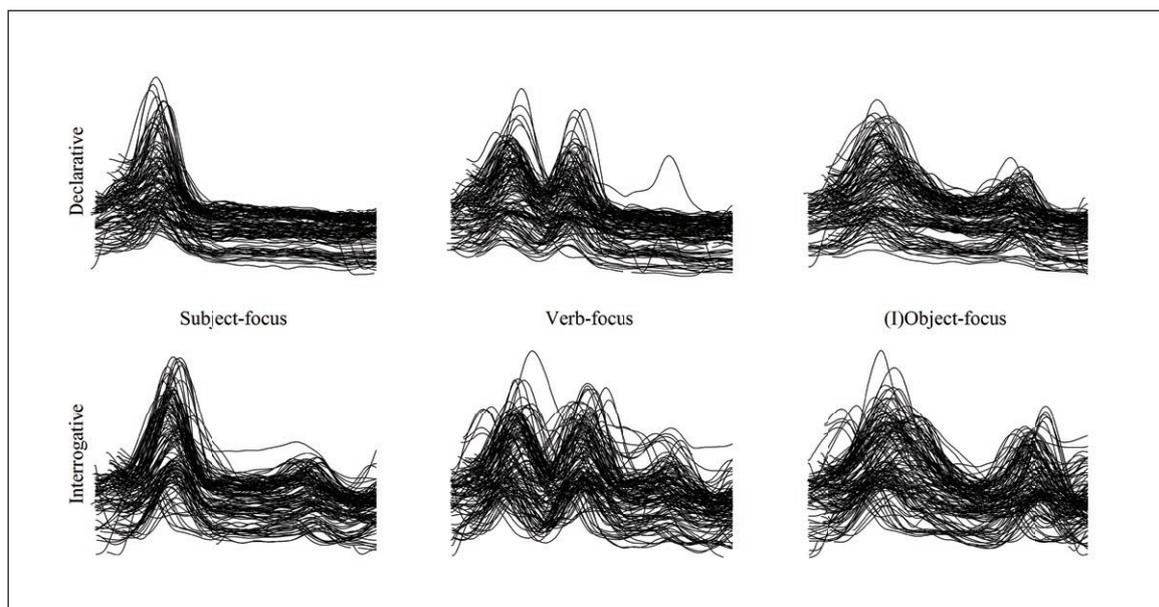
Figure 4: Utterance-long time normalised F_0 contours, in two sentence modalities (upper panels: declarative, lower panels: interrogative) and three focus positions (subject, verb, indirect object).

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Art. 9, page 7 of 20

Even if F_0 contours are sampled using the same number of points in time, variability might still surface as an artefact of differences in overall duration, although, overall utterance duration did not play a role in signalling sentence modality in a previous study of Neapolitan Italian (Cangemi & D'Imperio, 2011). The role of durational cues was investigated by predicting Utterance Duration with a mixed effects linear model featuring Modality (Declarative, Interrogative), Focus Placement (Subject, Verb, Object) and their interaction as fixed effects, Speaker, Sentence, and Repetition as random intercepts, and by-Speaker and by-Sentence random slopes for Modality, Focus, and their interaction. Neither Modality nor the interaction between Modality and Focus reached significance (all $|t| < 1.69$). Likelihood ratio tests between the full model and a simpler model (in which Modality and the interaction between Modality and Focus are dropped) show no significant difference (all $p > 0.25$). These results indicate that the different degree of variability found across sentence modalities is not a by-product of durational differences.
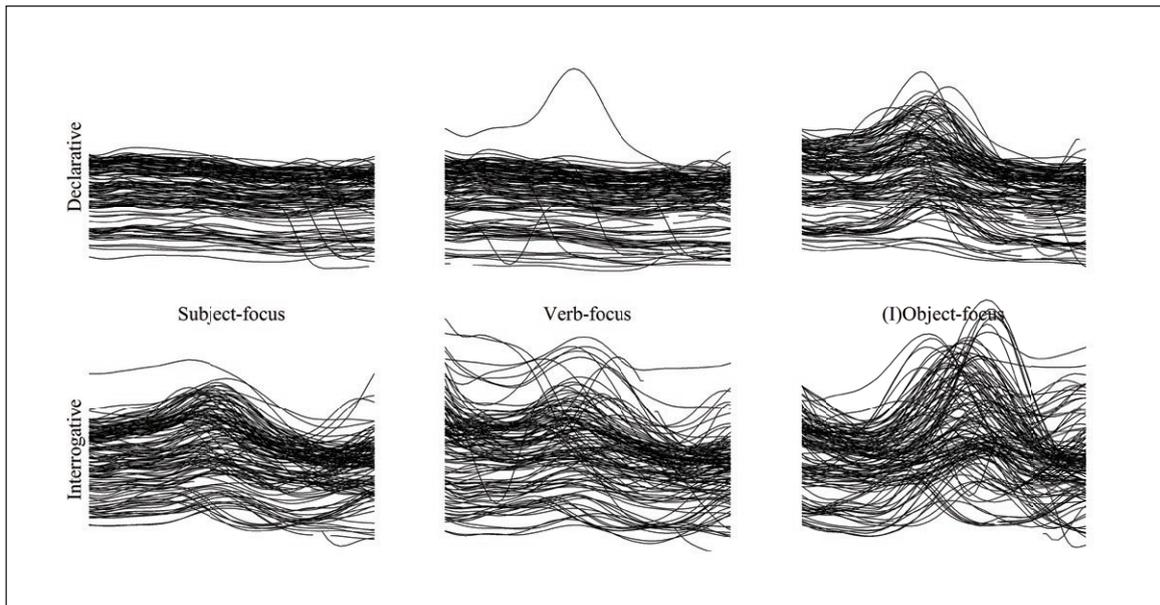
Figure 4 also illustrates the well-documented finding that in Neapolitan Italian, as in other varieties (e.g., Palermo (Grice, 1995) and Bari (Grice & Savino, 1997; Grice et al., 2005)), the final constituent has a pitch peak regardless of the location of focus in the sentence. This peak is usually reduced in range if the focus is earlier: "yes/no questions with early focus present [...] a smaller peak on the last stressed syllable of the intonational phrase" (D'Imperio, 1997, p. 25; see also D'Imperio, 2001). Obviously, peaks are intrinsically more prone to variation than flat stretches. It might thus be that variability is a mere by-product of the number of peaks in an utterance. In the subject- and verb-focus conditions, interrogatives (with two peaks) are thus expected to be more variable than declaratives (which have one). However, in the object-focus condition, both modalities have two peaks, and as we have seen above, interrogatives are still more variable than declaratives. This fact suggests that variability is not simply a matter of how many pitch movements are present. Under this assumption, we tested a more conservative prediction of our hypothesis, according to which interrogatives are more variable than declaratives even when the final prosodic word is analysed separately. Levene's tests support this prediction as well, both for utterances gated before the final prosodic word and, as shown in **Figure 5**, for the final prosodic word by itself (all $p < 0.01$).

2.3. Macroscopic analysis (F_0 across utterances): sub-clustering

The final prosodic word deserves particular attention, especially since (as we suggested above, Section 1.2) declaratives consistently show final falls, whereas interrogatives display either final rises or final falls. The greater variability in interrogatives might thus reflect either (i) the fact that one pragmatic category (interrogative) can be represented by two sub-clusters (final rise vs. final fall), or (ii) that dispersion of actual realizations is higher for interrogatives independently of differences in sub-clustering—or both.

In order to explore sub-clustering, we automatically classified utterance-final contours into rising and falling. Contours were classified on the basis of the difference between the mean F_0 in the final portion of the prosodic word (its last 10 samples) and the immediately preceding stretch (with the same duration). If the delta exceeded 10Hz, items were classified as rising.[2] **Figure 6** shows the results of the automatic classification, with rising contours plotted in red, for the two modalities (declarative, top panels; interrogative, bottom panels) and the three focus conditions (subject-focus, left panels; verb-focus, central

---

[2] We used an absolute threshold instead of empirically determined just noticeable differences since the listener-specific sensitivity thresholds reported by 't Hart (1981) could be even more relevant in utterance-final position.

Art. 9, page 8 of 20

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Figure 5: Fundamental frequency contours for final prosodic words. Vertical range 75–300 Hz. Declaratives (upper panels), interrogatives (lower), in three focus positions, subject focus (left), verb focus (middle), and indirect object focus (right).

panels; indirect object-focus, right panels). As expected, with only few exceptions (4 cases out of 123 subject-focus declaratives and 5 cases out of 120 verb-focus declaratives), final rises appear in interrogatives only.

The greater variability in interrogatives is not only due to this final rise, however: Levene's tests on items with final falls confirm that, even in this reduced dataset, interrogatives are realized more variably than declaratives (all $p < 0.001$). Thus, both sub-clustering and dispersion are responsible for the greater variability of F_0 contours in interrogatives:

(i) *Sub-clustering.* Interrogatives can be realized with two different strategies (either with a final fall or final rise), while declaratives are restricted to final falls.

(ii) *Dispersion.* Even when focussing on one single cluster (final falls), realizations are still more variable in interrogatives.

2.4 Microscopic analysis (peak alignment in pitch accents): sub-clustering and dispersion

So far, we have explored the interplay between sub-clustering and dispersion at a macroscopic level, by evaluating variance in F_0 values across entire utterances or prosodic words. In this paragraph, we show that these same two sources of variability also operate at a microscopic level. In order to do this, we will focus on peak alignment in the last (nuclear) pitch accent of object-focused utterances. Peak alignment has been shown to be an important cue in distinguishing declaratives from interrogatives in Neapolitan Italian. It is for this reason that we focus on this aspect of phonetic substance here.

Peak alignment was automatically extracted using a procedure in four steps. First, the F_0 contours of the last prosodic word were extracted using 50 equally spaced sampling points for each item, thus intrinsically normalizing for durational differences. Then we extracted the number of local maxima; some items ($n = 76$) had a single maximum, but most had more than one; only very few ($n = 13$) had more than four, and were discarded from analysis. Visual inspection of the remaining cases with two to four maxima showed that two items contained two non-adjacent samples with the same exact F_0 values; these were excluded from analysis. All other items had adjacent maxima, viz. very short plateaux (< 25ms). In such cases, the peak was located at the end of the plateau (D'Imperio, 2000;
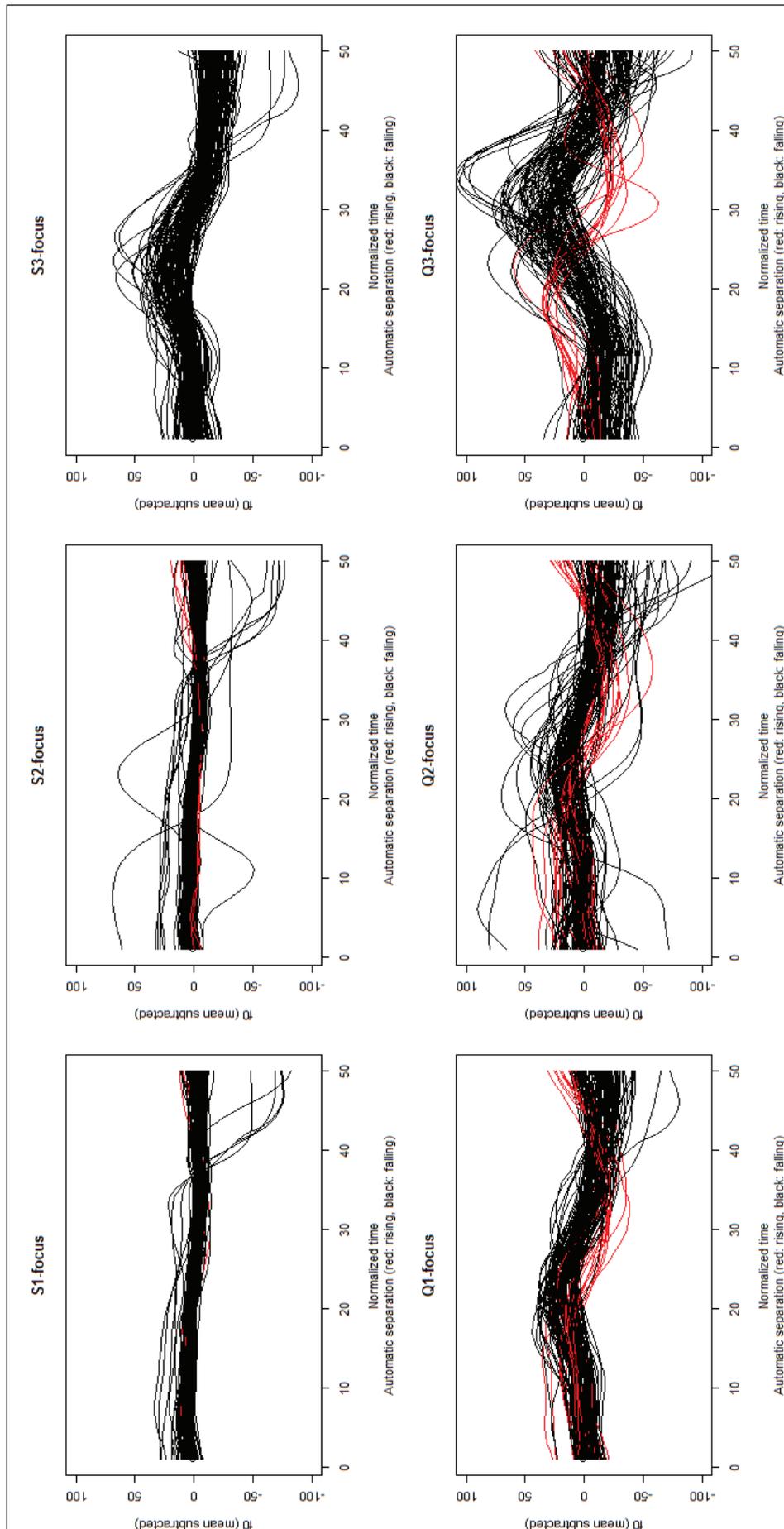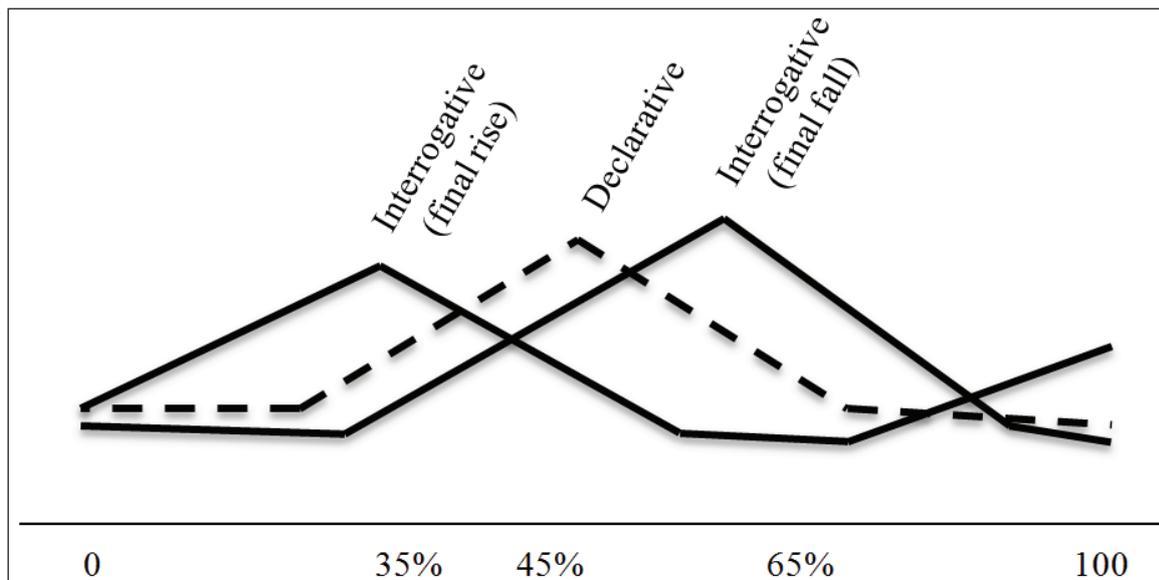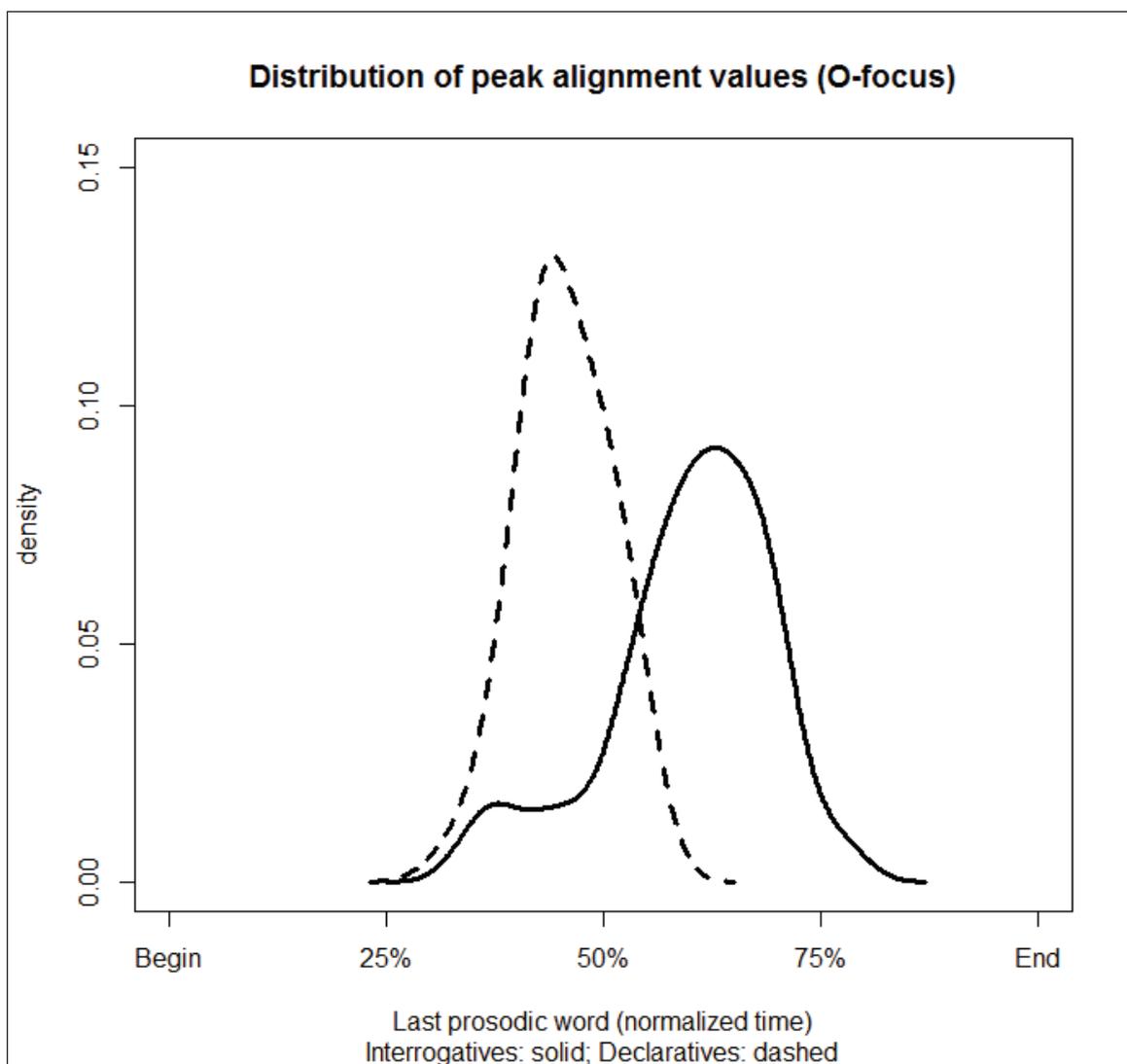
Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Art. 9, page 9 of 20

Figure 6: Automatic classification of final rises. F_0 values centered around the mean.

Art. 9, page 10 of 20

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Figure 7: Schematic pitch contours on the last prosodic word in declaratives (dashed line) and interrogatives (solid line).

Figure 8: Distribution of peak alignment for each sentence modality.

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Art. 9, page 11 of 20

Knight & Nolan, 2006). For six items the peaks located in this way surfaced in the first or final fifth of the last prosodic word (hence well away from the medial stressed syllable) and were discarded as artefacts. Since two items were discarded due to disfluencies, the final dataset resulted in 231 items. **Figure 7** shows a schematic representation of intonational contours for the two sentence modalities and the two edge tones.

Figure 8 shows the distribution of peak alignment values for declaratives (dashed line) and interrogatives (solid line). In many ways the results mirror those for the overall F_0 contours across entire utterances. First, interrogatives show sub-clustering, as can be seen by the bimodal distribution. The shoulder around 35% in normalized time of the last prosodic word is in fact the contribution of interrogatives with final rise, in which the peak is retracted in order to accommodate the following rise (see also **Figure 6**, right panels, and **Figure 7** for actual and schematic pitch contours respectively). Declaratives, on the other hand, show a single peak around 45%. Moreover, the second peak in interrogatives, situated around 65%, has a larger kurtosis than the peak for declaratives. This indicates that even when only final falls are taken into account, dispersion is still higher in interrogatives than declaratives. This patterning is confirmed by Levene's tests contrasting variability of peak alignment in declaratives and interrogatives, both when including ($p < 0.01$) and excluding ($p = 0.01$) interrogatives with final rises.

3 Discussion

3.1 Summary of results

The exploration of the Neapolitan Italian read speech corpus has shown that pitch contours are more variable in interrogatives than in declaratives. This is true both at a macroscopic level, i.e., in terms of variability of F_0 tracks across the entire utterance (see Sections 2.2–3), and at a microscopic level, i.e., in terms of variability of peak alignment within individual pitch accents (Section 2.4).

Macroscopic analysis. Interrogatives have been shown to be encoded more variably than declaratives in all portions of the utterance, independently of focus placement and of durational differences (Section 2.2). The variability in interrogatives has been ascribed to two sources (Section 2.3): the presence of sub-clusters (final rises or final falls), and a higher dispersion of values within a sub-cluster (within the interrogatives with final falls).

Microscopic analysis. In focus-final utterances, peak alignment in interrogatives has been shown to be more variable than in declaratives (Section 2.4). This has been ascribed to the two same sources of variability invoked in the macroscopic analysis: the distribution of peak alignment values not only shows two clusters (late and very early peaks), but the late-peak cluster also has a higher kurtosis (i.e., has more dispersed realizations).

In the following, we speculate on some possible causes and consequences of the greater variability in the encoding of interrogatives (Section 3.2). We conclude with a discussion on the implications of our findings towards the theory and practices of transcription, in particular prosodic transcription (Section 3.3).

3.2 On the sources and consequences of variability in interrogatives

In a distributional approach, one would of course expect variability across realizations of a given category. More importantly for our purposes, there is also no reason to assume that this degree of variability should be the same across different categories. One category might be instantiated by fairly variable tokens, while another category might be encoded more compactly. Our results show indeed that interrogatives are encoded more variably than declaratives. It is important to take a closer look at

Art. 9, page 12 of 20

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

these differences, since this state of affairs might emerge as a consequence of how categories are organized in a system (and thus provide insights on a language's prosodic system), and in turn be reflected in how such categories are built, used, and updated (and thus generate hypotheses on language acquisition, interaction, and sound change). An extensive discussion of the sources and consequences of such *differential variability* of categories in intonation is beyond the scope of this paper (for the notion of differential variability in a different domain, see Ho et al., 2008). In the following, we will provide only a brief overview of possible sources and one example of how the notion of differential variability might be useful in other research areas (viz. language contact). Implications for prosodic transcription will be dealt with more extensively in the final section (Section 3.3).

Escandell-Vidal (1998) explored the role of intonation in "procedural encoding" (Wilson & Sperber, 1993) of interrogatives, suggesting that speakers (in this case, of Peninsular Spanish) use different intonation contours in interrogatives in order to guide the listener towards a particular understanding of the content of an utterance. Recent experimental research suggests that intonation might encode a variety of pragmatic biases in interrogatives (Borràs-Comes & Prieto, 2014; Nilsenova, 2002; Savino, 2014a; Savino & Grice, 2011) across languages. These include epistemic (the assessment of the truth of a proposition, Sudo, 2013), evidential (availability of evidence, Büring & Gunlogson, 2000), mirative (surprise or unexpectedness, Peterson, 2010), and doxastic (disbelief, Crespo-Sendra et al., 2013) biases. At this point, it is unclear whether interrogatives are genuinely more prone than declaratives to an intonational encoding of pragmatic biases, or if such a picture is a mere consequence of the recent accumulation of experimental research focusing on interrogatives rather than declaratives. Similarly, one might ask whether intonation research has so far adopted an excessively broad understanding of interrogativity, thus somehow neglecting the development of specific methodological paradigms to effectively tell apart such "nuances". As mentioned earlier (Section 1.2), interrogatives have also been shown to be more variable in terms of regional variation, in particular across varieties of Italian (Savino, 2012). This might in turn motivate the differences found across speech styles (Grice et al., 1997), especially in diglossic contexts, or if speakers have access to a highly stratified repertoire. Similarly, politeness is also argued to play a role in the selection of specific intonation contours for interrogatives (Astruc et al., in press; Cruttenden, 1986).[3]

Given this picture, the notion of differential variability might prove useful in generating new research hypotheses and in accounting for some recent findings. Studies on prosodic accommodation in overt (D'Imperio et al., 2014; D'Imperio & Sneed German, 2015) and covert (Savino, 2014b) imitation tasks show that speakers of Italian varieties can adapt their pitch accent and boundary tone choices in the production of interrogatives. Crucially, Romera and Elordieta (2013) report that prosodic accommodation due to language contact is stronger for interrogatives than for declaratives. They analyze intonation patterns from a corpus of symmetrical semi-directed sociolinguistic interviews, with a native speaker of Majorcan Catalan (also an L2 speaker of the Majorcan variety of Spanish) as interviewer. The subjects were four speakers with monolingual Peninsular Spanish origins, who had been living in Majorca for 5 years on average at the time of the interview. Whereas no subject used Majorcan Spanish intonation in their production of declaratives, all subjects used Majorcan Spanish intonation for interrogatives in at

---

[3] As in the case of pragmatic biases, we do not exclude that politeness might affect declaratives in equal measure. It may therefore be an issue of what happens to pitch accents when followed by a final rise instead of a fall, rather than interrogatives per se.

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Art. 9, page 13 of 20

least 25% of their productions. Along the lines of Trudgill (1986), the authors suggest that speakers might accommodate their production of interrogatives more readily to the Majorcan Spanish prototype because these are perceptually more salient, and thus considered to be more characteristic of the target variety. However, if interrogatives were encoded more variably in Peninsular Spanish, the higher degree of accommodation of interrogatives to the Majorcan Spanish target might also stem from the internal structure of the interrogative category itself. Accommodation might be easier if the source category has a rich structure, with sub-clusters and high internal dispersion (as with interrogatives), rather than a very peaked distribution (as with declaratives), resulting in more entrenched productions.[4]

3.3 Implications for intonation transcription

The alignment of F_0 peaks is an important aspect of phonetic substance that is taken into account when deciding on an inventory of intonational categories, and later when deciding on category membership, in particular pitch accent type. Extensive research on categorical perception has focused on this cue to the interrogative-declarative distinction (D'Imperio & House, 1997). Nonetheless, even when investigating this very distinction, we have identified tonal contexts that can radically affect the alignment of F_0 peaks. Taking only the phonetic substance into account, pitch accents in interrogatives with a final rise should pattern with a pitch accent type involving a medial peak (L + H*, first analysed as H* + L, see D'Imperio, 2001; Grice et al., 2005; and Frota, 2016, for discussion). However, taking meaning into account, they should pattern with pitch accents with a late peak (L* + H). At first sight, this might look like a good argument for a broad phonetic level of transcription dealing with substance, and a phonological level of transcription dealing with meaning. However, it is unclear whether we actually hear these peaks as medial, as other cues might be at work, especially since it is not clear how modular our perception of pitch accents and following boundary tones is (Dainora, 2006). It is thus unclear whether we integrate the position of the peak with (possibly language-specific) adjustments made owing to tonal context. In this case the anticipation of the peak would serve to ensure that the tonal contour is realized, despite the lack of sufficient segmental material to bear the tones assigned to it.

An intonation transcription system needs to have mechanisms for dealing with contextually determined variation, i.e., adjustments due to tonal crowding. Adjustments can be made to the articulation rate: slowing down facilitates the accommodation of the tones (Erikson & Alstermark, 1972). However, rate adjustments do not involve a uniform lengthening of segments (Byrd & Saltzman, 2003). Thus, the alignment of intonational peaks will involve more or less dispersion across conditions, depending on the landmarks selected for reference. Alternatively, or in addition, the tones may move closer together, leading to faster pitch movements. In this study a different kind of adjustment was apparent: The whole tonal sequence starts earlier in relation to the segmental string. Anticipation has been found in Tashlhiyt Berber (Grice et al., 2015), in cases where the segmental string was entirely voiceless and thus not tone-bearing. It has also been found in Dutch (Hanssen et al., 2007) in cases of tonal crowding. All of these sources of variation make absolute and relative alignment measures difficult to interpret.

Another source of variation is truncation, a process in which tones undershoot their targets. Naturally, the transcriber is faced with the decision as to whether a tone is there but only partially realised, i.e., truncated, or simply not there at all. Take, for instance,

---

[4] Note that *declaratives* refers here to contrastive statements, as opposed to contrastive questions (*interrogatives*). Other focus structures, such as non-contrastive broad focus, might be characterized by greater variability.

Art. 9, page 14 of 20

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

contours that Grabe (2008) analyzed as a truncated fall in German. The decision to refer to them as truncated *falls* might appear to have been made on the basis of the meaning of a given contour, rather than its phonetic substance (in this case the F_0 trace). Thus, a high short level tone on <Schiff> may be interpreted as a truncated fall by virtue of its functional equivalence to <Schiefer> in a declarative sentence, which has a clear fall over the two syllables (Grabe, 2008). However, recent work has shown that the spectral characteristics of voiceless segments can give an impression of pitch (Kohler, 2011; Niebuhr, 2008; Ritter & Roettger, 2014), so that the phonetic substance—in the form of perceptual integration of cues—most probably played a role in the categorization, despite the lack of movement in the fundamental frequency contour. This indicates that multiple cues make it difficult to decide on the basis of a selective visual representation of the signal—especially when this involves a visual representation of F_0—as to category membership. This is to be expected, given that the cues encoding speech in general are acoustically diverse and functionally redundant (see Winter, 2014).

The data presented in this study raise the question as to whether peak alignment is a suitable cue in itself. In fact, it has been treated as an abstraction by Gussenhoven (2004), who shows that peak delay can be used as an enhancement of—or substitute for—peak raising. Tradeoffs between peak alignment and other parameters such as peak scaling or shape have been documented for Russian (Rathcke, 2006) and German (Ambrazaitis & Frid, 2014; Niebuhr, 2007). Furthermore, Niebuhr et al. (2011) show that there are considerable individual differences in the implementation of intonational categories, specifically, that for some speakers the *shape* of the F_0 trajectory may be used, instead of the absolute *alignment* of a peak, to signal the same category. Therefore we should exercise caution when reducing contrasts to one dimension.

Let us examine a typical contrast in the segmental domain that is frequently discussed as analogous to peak alignment, the distinction between 'voiceless' and 'voiced' oral stops, e.g., between /p/ and /b/ (see also Arvaniti, 2016). This distinction is frequently invoked in the intonation literature in relation to one dimension—voice onset time (VOT). VOT shares with F_0 peak alignment the crucial timing of glottal and supralaryngeal gestures, although VOT is concerned with onset of vibration of the vocal folds, whereas peak alignment is concerned with modulating the frequency of vibration.

In many accounts, the terms *fortis* and *lenis* are used for this distinction, rather than *voiceless* and *voiced,* reflecting the fact that voicing (vocal fold vibration) is not the only cue involved in the contrast. Comparing fortis and lenis plosives acoustically, fortis plosives have a longer closure duration, and a longer and stronger burst, resulting from a longer articulatory constriction duration and a higher intraoral peak pressure. Moreover, in intervocalic position the preceding vowel is longer and the transitions into the vowel are more abrupt (Kohler, 1979; Lisker, 1978; Slis & Cohen, 1969). Despite the term *micro*-prosody, the fortis-lenis distinction can have a considerable effect on the F_0, and can be used as a cue to voicing, even when voice onset time cues are unambiguous (see, e.g., Kingston, 2007; Whalen et al., 1993).

Furthermore, despite the emphasis in the literature on VOT, the aspiration (i.e., positive VOT) in English is often drastically reduced in weak prosodic positions (such as in the unstressed syllable in 'rapid' or word finally in 'hip') or after a word initial sibilant (in 'spin'). In these cases it is unlikely that a lenis symbol is selected, as the transcriber is aware of the contextually determined variation, and of course keeps the lexical meaning in mind.

This is less obviously the case when transcribing obstruents across dialects. Barry and Pützer (1995) point out that transcribers might weight the different cues to the fortis-lenis distinction in different ways, leading to different choices of symbol, even for the same

Cangemi and Grice: The Importance of a Distributional Approach to Categoriality in Autosegmental-Metrical Accounts of Intonation

Art. 9, page 15 of 20

dialect. Their study looks at cognates in Mosel-franconian and Rhenish-franconian. Mosel-franconian /tan/ (Standard German. *Tanne*, Eng. 'fir tree') has a cognate in Rhenish-franconian dialects that is transcribed in a number of pronunciation dictionaries as /dan/. However, this poses a system-internal problem, as there are minimal pairs in Rhenish-franconian, e.g., /pE:r/ /bE:r/ (horse-bear). Specifically, they found in this study that the duration and strength of the burst was the most important cue for native listeners. They hence argue for using the fortis symbol for the cognate of /tan/ in Rhenish-franconian on the grounds that the plosive has fortis cues (burst properties, i.e. form-substance), even if one cue, aspiration (i.e., VOT), was absent. This is also in accord with the production of fortis plosives as contrasted with lenis plosives (minimal pairs, i.e., form-meaning).

Thus, even in the so-called segmental domain there are problems with categorization. Just as /p/ is selected by the transcriber to represent the sound in "spin" despite its zero VOT, L*+H might be selected by the transcriber in the rise-fall-rise case, despite the early alignment of the F_0 peak. In both cases the meaning and knowledge of context-dependent variation would guide the choice. If instead, the transcriber relied purely on phonetic substance, in both cases on one cue dimension (VOT and peak alignment), then /b/ and L+H* might be selected. If both meaning and substance were taken into account, but with more attention to phonetic detail, then [p] would be selected with a diacritic (or in this case absence of a diacritic for aspiration). But the conventions for transcribing /p/ in unstressed syllables ("rapid") and coda position ("hip") are less straightforward. For instance, in German narrow phonetic transcription it is customary to transcribe aspiration in word final plosives, whereas this is not the case for English, aspiration being transcribed in prevocalic position. The level of granularity used in transcription naturally depends on the purpose (Ladefoged, 1990), but the question remains as to what is behind the decision to opt for one symbol over another (reflecting category membership at the level of form). This question holds for intontional categories as well.

Supplementary Files

For accompanying TextGrid and wav files, go to http://dx.doi.org/10.5334/labphon.28.smo

Acknowledgements

Competing Interests

The authors declare that they have no competing interests.

References

Ambrazaitis, G., & Frid, J. 2014. F0 peak timing, height, and shape as independent features. *Proceedings of the 4th International Symposium on Tonal Aspects of Languages, (Nijmegen)*, 138–142.

André, C., Ghio, A., Cavé, C., & Teston, B. 2003. Perceval: a computer-driven system for experimentation on auditory and visual perception. *Proceedings of the 15th International Congress of Phonetic Sciences (Barcelona)*, pp. 1421–1424.

Arvaniti, A. 2016. Analytical Decisions in Intonation Research and the Role of Representations: Lessons from Romani. *Laboratory Phonology: Journal of the Association for Laboratory Phonology,* 7(1): 6, pp. 1–43, DOI: http://dx.doi.org/10.5334/labphon.14

Art. 9, page 16 of 20

Cangemi and Grice: The Importance of a Distributional Approach to Categoriality in Autosegmental-Metrical Accounts of Intonation

Astruc, L., Vanrell, M. M., & Prieto, P. in press. Cost of the action and social distance affect the selection of question intonation in Catalan. In Armstrong, M., Henriksen, N., & Vanrell, M. M. (Eds.), *Interdisciplinary approaches to intonational grammar in Ibero-Romance intonation*. Amsterdam: John Benjamins.

Audacity Development Team. 2006. *Audacity: free audio editor and recorder*. Computer program.

Barry, W. J., & Pützer, M. 1995. Zur phonetischen Basis der Fortis-Lenis-Opposition bei Plosiven in moselfränkischen und rheinfränkischen Dialektgebieten im Saarland und in Rheinland-Pfalz. *Phonus, 1*, 53–64.

Bates D., Maechler, M., Bolker, B., & Walker, S. 2014. *lme4: Linear mixed-effects models using Eigen and S4*. Computer program.

Baumann, S., Grice, M., & Steindamm, S. 2006. Prosodic marking of focus domains – categorical or gradient? *Proceedings of Speech Prosody 2006 (Dresden)*, pp. 301–304.

Boersma, P., & Weenink, D. 2008. *Praat: doing phonetics by computer*. Computer program.

Borràs-Comes, J., & Prieto, P. 2014. Intonational meaning and the pragmatics of yes-no questions in Central Catalan. *Talk at the 7th Conference on Laboratory Approaches to Romance Phonology (Aix-en-Provence)*.

Büring, D., & Gunlogson, C. 2000. Aren't positive and negative polar questions the same? Unpublished manuscript. Available at http://hdl.handle.net/1802/1432.

Byrd, D., & Saltzman, E. 2003. The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics, 31*, 149–180. DOI: http://dx.doi.org/10.1016/S0095-4470(02)00085-2

Cangemi, F. 2014. *Prosodic detail in Neapolitan Italian*. Berlin: Language Science Press.

Cangemi, F., Cutugno, F., Ludusan, B., Seppi, D., & Van Compernolle, D. 2011. ASSI – Automatic Speech Segmentation for Italian: tools models, evaluation and applications. In Gili Fivela, B., Stella, A., Garrapa, L., & Grimaldi, M. (Eds.), *Contesto comunicativo e variabilità nella produzione e percezione della lingua, (Atti del VII Convegno Nazionale AISV, Lecce)*. Roma: Bulzoni Editore, pp. 337–344.

Cangemi, F., & D'Imperio, M. 2011. Tempo, focus and modality in Neapolitan Italian. *Proceedings of the 17th International Congress of Phonetic Sciences (Hong Kong)*, pp. 392–395.

Cangemi, F., & D'Imperio, M. 2013. Tempo and the perception of sentence modality. *Journal of the Association for Laboratory Phonology, 4*(1), 191–219. DOI: http://dx.doi.org/10.1515/lp-2013-0008

Cole, J., & Shattuck-Hufnagel, S. 2016. New Methods for Prosodic Transcription: Capturing Variability as a Source of Information. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1): 8, pp. 1–29, DOI: http://dx.doi.org/10.5334/labphon.29

Coleman, J. 2003. Discovering the acoustic correlates of phonological contrasts. *Journal of Phonetics, 31*, 351–372. DOI: http://dx.doi.org/10.1016/j.wocn.2003.10.001

Crespo-Sendra, V., Kaland, C., Swerts, M., & Prieto, P. 2013. Perceiving incredulity: The role of intonation and facial gestures. *Journal of Pragmatics, 47*(1), 1–13. DOI: http://dx.doi.org/10.1016/j.pragma.2012.08.008

Cruttenden, A. 1986. *Intonation*. Cambridge: Cambridge University Press.

Crystal, D. 1975. Intonation and linguistic theory. In Dahlstedt, K-H. (Ed.), *The Nordic languages and modern linguistics 2*. Stockholm: Almqvist & Wiksell, pp. 267–303.

Dainora, A. 2006. Modeling intonation in English: A probabilistic approach to phonological competence. In Goldstein, L., Whalen, D., & Best, C. (Eds.), *Laboratory Phonology 8*. Berlin, New York: Mouton de Gruyter, pp. 107–132.

de Saussure, F. 1916. Nature of the linguistic sign. In Bally, C., & Sechehaye, A. (Eds.), *Cours de linguistique générale*. New York: McGraw Hill Education.

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Art. 9, page 17 of 20

D'Imperio, M. 1997. Breadth of focus, modality and prominence perception in Neapolitan Italian. *Ohio State University Working Papers in Linguistics, 50*, 19–39.

D'Imperio, M. 2000. *The role of perception in defining tonal targets and their alignment* (Unpublished doctoral dissertation). The Ohio State University. DOI: http://dx.doi.org/10.1016/S0167-6393(00)00064-9

D'Imperio, M. 2001. Focus and tonal structure in Neapolitan Italian. *Speech Communication, 33*(4), 339–356.

D'Imperio, M. 2002. Italian intonation: an overview and some questions. *Probus, 14*(1), 37–69. DOI: http://dx.doi.org/10.1515/prbs.2002.005

D'Imperio, M., Cavone, R., & Petrone, C. 2014. Phonetic and phonological imitation of intonation in two varieties of Italian. *Frontiers in Psychology, 5*, 1226. DOI: http://dx.doi.org/10.3389/fpsyg.2014.01226

D'Imperio, M., & House, D. 1997. Perception of questions and statements in Neapolitan Italian. *Proceedings of Eurospeech, 1*, 251–254.

D'Imperio, M., & Sneed German, J. 2015. Phonetic detail and the role of exposure in dialect imitation. *Proceedings of the 18th International Congress of Phonetic Sciences (Glasgow)*.

Erikson, Y., & Alstermark, M. 1972. Fundamental frequency correlates of the grave word accent in Swedish: the effect of vowel duration. *Speech Transmission Laboratory, Quarterly Progress and Status Report*, 2–3. KTH, Sweden.

Escandell-Vidal, V. 1998. Intonation and procedural encoding: the case of Spanish interrogatives. In Rouchota, V., & Jucker, A. (Eds.), *Current issues in relevance theory*. Amsterdam: John Benjamins, pp. 169–203. DOI: http://dx.doi.org/10.1075/pbns.58.09esc

Fox, J., & Weisberg, S. 2011. *An R companion to applied regression*. Thousand Oaks, CA: Sage.

Frota, S. 2016. Surface and Structure: Transcribing Intonation within and across Languages. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1): 7, pp. 1–19. DOI: http://dx.doi.org/10.5334/labphon.10

Gafos, A., & Beňuš, S. 2006. Dynamics of phonological cognition. *Cognitive Science, 30*(5), 905–943. DOI: http://dx.doi.org/10.1207/s15516709cog0000_80

Gili Fivela, B., Avesani, C., Barone, M., Bocci, G., Crocco, C., D'Imperio, M., Giordano, R., Marotta, G., Savino, M., & Sorianello, P. 2015. Varieties of Italian and their intonational phonology. In Frota, S., & Prieto, P. (Eds.), *Intonational variation in Romance*. Oxford: Oxford University Press, pp. 140–197. DOI: http://dx.doi.org/10.1093/acprof:oso/9780199685332.003.0005

Grabe, E. 2008. Pitch accent realisation in English and German. *Journal of Phonetics, 26*(2), 129–143. DOI: http://dx.doi.org/10.1006/jpho.1997.0072

Grice, M. 1995. *The intonation of interrogation in Palermo Italian – Implications for intonation theory*. Tübingen: Niemeyer.

Grice, M., D'Imperio, M., Savino, M., & Avesani, C. 2005. Strategies for intonation labelling across varieties of Italian. In Jun, S.-A., (Ed.), *Prosodic typology*. Oxford: Oxford University Press, pp. 362–389. DOI: http://dx.doi.org/10.1093/acprof:oso/9780199249633.003.0013

Grice, M., Ridouane, R., & Roettger, T. B. 2015. Tonal association in Tashlhiyt Berber. Evidence from polar questions and contrastive statements. *Phonology*, *32*(2), 241–266.

Grice, M., & Savino, M. 1997. Can pitch accent type convey information status in yes-no questions? *Proceedings of the ACL workshop "Concept to Speech Generation Systems" (Madrid)*, pp. 29–38.

Grice, M., & Savino, M. 2003. Map Tasks in Italian: asking questions about given, accessible and new information. *Catalan Journal of Linguistics, 2*, 153–180.

Art. 9, page 18 of 20

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Grice, M., Savino, M., & Refice, M. 1997. The intonation of questions in Italian: do speakers replicate their spontaneous speech when reading? *Phonus, 3*, 1–7.

Gussenhoven, C. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press. DOI: http://dx.doi.org/10.1017/CBO9780511616983

Halliday, M. 1967. *Intonation and grammar in British English*. The Hague: Mouton. DOI: http://dx.doi.org/10.1515/9783111357447

Hanssen, J., Peters, J., & Gussenhoven, C. 2007. Phrase-final pitch accommodation effects in Dutch. *Proceedings of the 16th International Congress of Phonetic Sciences, (Saarbrücken)*, pp. 1077–1080.

Ho, J., Stefani, M., dos Remedios, C., & Charleston, M. 2008. Differential variability analysis of gene expression and its application to human diseases. *Bioinformatics, 24*(13), 390–398. DOI: http://dx.doi.org/10.1093/bioinformatics/btn142

House, D. 2005. Phrase-final rises as a prosodic feature in wh-questions in Swedish human–machine dialogue. *Speech Communication, 46*, 268–283. DOI: http://dx.doi.org/10.1016/j.specom.2005.03.009

Jun, S.-A. (Ed.) 2005. *Prosodic typology*. Oxford: Oxford University Press. DOI: http://dx.doi.org/10.1093/acprof:oso/9780199249633.001.0001

Kingston, J. 2007. F0 Segmental influences on F0: automatic or controlled? In Gussenhoven, C., & Riad, T. (Eds.), *Tones and Tunes vol. 2*. Berlin: Mouton de Gruyter, pp. 171–210. DOI: http://dx.doi.org/10.1515/9783110207576.2.171

Knight, R.-A., & Nolan, F. 2006. The effect of pitch span on intonational plateaux. *Journal of the International Phonetic Association, 36*(1), 21–38. DOI: http://dx.doi.org/10.1017/S0025100306002349

Kohler, K. 1979. Dimensions in the perception of fortis and lenis plosives. *Phonetica, 36*, 332–343. DOI: http://dx.doi.org/10.1159/000259970

Kohler, K. 2004. Pragmatic and attitudinal meanings of pitch patterns in German syntactically marked questions. *Working Papers of the Kiel Institute of Phonetics and Digital Speech Processing (AIPUK),* 35a, pp. 125–142.

Kohler, K. 2011. Communicative functions integrate segments in prosodies and prosodies in segments. *Phonetica, 68*, 26–56. DOI: http://dx.doi.org/10.1159/000328433

Kügler, F. 2003. Do we know the answer? Variation in yes-no-question intonation. In Fischer, S.,van de Vijver, R., & Vogel, R. (Eds.), *Experimental studies in linguistics,* 1, 9–29.

Ladd, D. R. 2014. *Simultaneous structure in phonology*. Oxford: Oxford University Press. DOI: http://dx.doi.org/10.1093/acprof:oso/9780199670970.001.0001

Ladefoged, P. 1990. The revised International Phonetic Alphabet. *Language, 66*, 550–552. DOI: http://dx.doi.org/10.2307/414611

Lisker, L. 1978. Rapid vs. rabid: A catalogue of acoustic features that may cue the distinction. *Haskins Laboratories: Status Report on Speech Research* SR-54.

Niebuhr, O. 2007. The signalling of German rising-falling intonation categories – The interplay of synchronization, shape, and height. *Phonetica, 64*, 174–193. DOI: http://dx.doi.org/10.1159/000107915

Niebuhr, O. 2008. Coding of intonational meanings beyond F0: evidence from utterance-final /t/ aspiration in German. *Journal of the Acoustical Society of America, 124*, 1252–1263. DOI: http://dx.doi.org/10.1121/1.2940588

Niebuhr, O., D'Imperio, M., Gili Fivela, B., & Cangemi, F. 2011. Are there "shapers" and "aligners"? Individual differences in signalling pitch accent category. *Proceedings of the 17th International Congress on Phonetic Sciences (Hong Kong)*, pp. 120–123.

Nilsenova, M. 2002. A game-theoretical approach to the meaning of intonation in rising declaratives and negative polar questions. *Proceedings of the International Conference on Speech Prosody (Aix-en-Provence)*.

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Art. 9, page 19 of 20

Nolan, F. 2008. Intonation. In Aarts, B., & McMahon, A. (Eds.), *The handbook of English linguistics*, §19. Malden: Blackwell Publishing.

Ohala, J. 1990. There is no interface between phonetics and phonology. *Journal of Phonetics, 18*, 153–171.

Peterson, T. 2010. Examining the mirative and nonliteral uses of evidentials. In Peterson, T., & Sauerland, U. (Eds.), *Evidence from evidentials*. Working Papers in Linguistics. Vancouver: University of British Columbia, pp. 129–159.

Pierrehumbert, J., Beckman, M., & Ladd, D. R. 2000. Conceptual foundations of phonology as a laboratory science. In Burton-Roberts, N., Carr, P., & Docherty, G. (Eds.), *Phonological knowledge*. Oxford: Oxford University Press, pp. 273–303.

Rathcke, T. 2006. Relevance of f0-peak shape and alignment for the perception of a functional contrast in Russian. *Proceedings of Speech Prosody 2006 (Dresden)*, pp. 65–68.

R Development Core Team. 2008. *R: a language and environment for statistical computing.* Computer program.

Ritter, S., & Roettger, T. 2014. Speakers modulate noise-induced pitch according to intonational context. *Proceedings of the 7th International Conference on Speech Prosody (Dublin)*, pp. 890–893.

Romera, M., & Elordieta, G. 2013. Prosodic accommodation in language contact: Spanish intonation in Majorca. *International Journal of the Sociology of Language, 221*, 127–151. DOI: http://dx.doi.org/10.1515/ijsl-2013-0026

Savino, M. 2012. The intonation of polar questions in Italian: Where is the rise? *Journal of the International Phonetic Association, 42*(1), 23–48. DOI: http://dx.doi.org/10.1017/S002510031100048X

Savino, M. 2014a. Degrees of (un)certainty in Bari Italian yes-no question intonation: perceptual evidence. In Zuczkowski, A., Bongelli, R., Riccioni, I., & Canestrari, C., (Eds.), *The communication of certainty and uncertainty: Linguistic, psychological, philosophical aspects*. IADA online series, vol. 1/12, pp. 52–67.

Savino, M. 2014b. Prosodic adaptation in game interaction between speakers of two Italian varieties. *Proceedings of the 10th International Seminar on Speech Production (Cologne)*, pp. 371–374.

Savino, M., & Grice, M. 2011. The perception of negative bias in Bari Italian questions. In Frota, S., Elodierta, G., & Prieto, P. (Eds.), *Prosodic categories: production, perception and comprehension*. Dordrecht: Springer, pp. 187–206. DOI: http://dx.doi.org/10.1007/978-94-007-0137-3_8

Schafer, A., Carlson, K., Clifton, C., & Frazier, L. (2000). Focus and the interpretation of pitch accent. *Language and speech, 43*(1), 75–105. DOI: http://dx.doi.org/10.1177/00238309000430010301

Slis, I., & Cohen, A. (1969). On the complex regulating the voiced-voiceless distinction I and II. *Language and Speech, 12*, 80–102, 137–155.

Sudo, Y. (2013). Biased polar questions in English and Japanese. In Gutzmann, D., & Gaertner, H.-M. (Eds.), *Beyond expressives: Explorations in use-conditional meaning*. Leiden: Brill, pp. 275–296. DOI: http://dx.doi.org/10.1163/9789004183988_009

't Hart, J. 1981. Differential sensitivity to pitch distance, particularly in speech. *Journal of the Acoustical Society of America, 69*(3), 811–821. DOI: http://dx.doi.org/10.1121/1.385592

Trudgill, P. 1986. *Dialects in contact*. Oxford: Basil Blackwell.

Whalen, D., Abramson, A., Lisker, L., & Mody, M. 1993. F0 gives voicing information even with unambiguous voice onset times. *Journal of the Acoustical Society of America, 93*(4). DOI: http://dx.doi.org/10.1121/1.406678

Art. 9, page 20 of 20

Cangemi and Grice: The Importance of a Distributional Approach to
Categoriality in Autosegmental-Metrical Accounts of Intonation

Wilson, D., & Sperber, D. 1993. Linguistic form and relevance. *Lingua, 93*, 1–25. DOI: http://dx.doi.org/10.1016/0024-3841(93)90058-5

Winter, B. 2014. Spoken language achieves robustness and evolvability by exploiting degeneracy and neutrality. *BioEssays, 36*, 960–967. DOI: http://dx.doi.org/10.1002/bies.201400028