

Broad, narrow and contrastive focus in Florentine Italian

Cinzia Avesani* and Mario Vayra**

*ISTC-CNR, Padova, Italy

** Università di Siena ad Arezzo, Arezzo, Italy
avesani@csrf.pd.cnr.it, vayra@unisi.it

ABSTRACT

It is well established that focus may have prosodic reflexes in various languages. Previous data on Florentine Italian showed that broad focus and late narrow-contrastive focus utterances are marked by different pitch accents. With the present experiment we address the question whether a three way contrast exists in the intonational realization of broad, narrow-semantic and narrow-contrastive focus. Results show that while focus type (contrastive vs. non-contrastive) is signalled by different pitch accents, differences in focus scope (broad vs. narrow) are not.

1. INTRODUCTION

It is assumed that “focus” is a syntactic (or morphological) constituent which is marked in the surface structure and that foci are reflected both in semantic and pragmatic effects and in prosodic patterns. There are at least three linguistically relevant senses of ‘focus’ that can be distinguished in the literature: psychological, semantic and contrastive focus. The first corresponds to the psychological notion of *focus of attention* and refers to the attention state of speech participants in relation to some entity in the discourse (and it will not be treated here). The other two both refer to that part of a sentence that has been given prosodic prominence by the speaker and have different semantic and pragmatic effects. i) A constituent may represent the new information that has been asserted or questioned in relation to what has been called topic, presupposition, background. This is what we will refer to as “semantic focus”. This meaning of focus reflects the way in which the informational content of an event expressed by a sentence is represented and how its truth value is to be assessed. ii) A constituent (either a semantic focus or a topic) may also represent information that is being contrasted explicitly or implicitly with a limited set of alternatives. We will refer to this as “contrastive focus”.

An additional distinction often assumed in the intonational literature that we will consider here refers to focus scope. Typically, if the focal domain extends over whole sentences or whole constituents we have cases of “broad focus” (BF). “Narrow focus” (NF) in the literature refers to a focal domain which extends over individual words/constituents [e.g.:1:161]. A focus narrow in scope can be either of a semantic or of a contrastive type.

Focused constituents may be cued prosodically by prosodic phrasing, phrasal prominence or a combination of both. It has been assumed that the kind of prosodic

prominence which serves to mark focused constituents is the pitch accent. Pitch accent placement is constrained in such a way that the position of nuclear pitch accents in a given sentence can help in identifying the focused constituent(s) in languages like English, German, Dutch, Spanish, European Portuguese, Italian. If the nuclear accent is sentence final, though, it has long been claimed that ambiguity will arise in the interpretation of the scope of the focused constituent.

As for contrastive meaning, some authors claim that it is cued by specific pitch accents. They differ from those that convey newness either because are of a different type [e.g.: 2, 3] or because are more prominent; others claim that no such difference exists [5].

In all the varieties of Italian analysed so far, it has been shown that broad focus (BF) and narrow contrastive focus (NCF) are signalled by two different pitch accents, respectively falling and rising. BF pitch accents have been analysed as H+L* in all varieties. NCF pitch accents have been analysed as monotonal H* in Florentine, and bitonal H* accents with a leading or a trailing tone respectively in Neapolitan (L+H*) and Palermo and Bari Italian (H*+L) [5]. As a consequence, even if NCF occurs sentence finally, no ambiguity arises as far as broad/narrow scope reading [6]. Italian patterns in this respect with Spanish Portuguese [7] and crucially differs from German and Dutch [8], languages in which it is the placement of the nuclear accent, (together with absence or reduction of prefocal accents) and not the type of pitch accent selected that cues different focus readings.

Unfortunately, previous data on the intonation of focus structures in different varieties of Italian do not allow to distinguish the prosodic reflexes of contrastive vs. non-contrastive (semantic) narrow focus. Only for Bari Italian it has been reported that non-contrastive narrow focus declaratives can be produced with the same pitch accent as broad focus declaratives (H+L*), thus leading to the same ambiguity in the scope of focus reported for English and other languages [5].

Our own data on Florentine [9] allowed to assess a systematic difference in the marking of BF and NCF only. In both cases, the sentence-final nuclear syllable was characterized by a fall. By varying the number of postnuclear syllables, it was showed that the timing of the anchoring of H and L targets of the fall to the segmental string systematically differed in BF and NCF: L consistently aligned with the end of the word in NCF,

whereas it consistently aligned with the stressed syllable in BF. The data allowed us to characterise L as an edge-related tone in NCF and as an accentual tone in BF. The resulting analysis of two nuclear contours was: H+L*L-L% for BF and H*L-L% for NCF.

A recent proposal suggests that NSF and NCF have different phonological representations [10]. The authors claim that the prosodic tree of a NCF sentence, but not of a NSF one, undergoes restructuring processes that affect its *phrasing*. If the contrast docks on the last node of a phonological phrase, an intonational phrase boundary is inserted immediately after it. However, it is clear that the phrasing effects emerges non-ambiguously only if the contrast docks on a constituent that occurs sentence-medially: if the constituent is the last one in a phonological phrase that is sentence-final, it is already followed by an intonational phrase boundary.

In this paper we address the question whether a three way contrast exists in the intonational realization of the different focus types. We have analysed broad focus, late narrow-semantic and late narrow-contrastive focus declaratives. In all cases, the nuclear accent occurs sentence finally. In such a position the restructuring effects claimed by [10] to affect phrasing in NCF are neutralized. Hence, if a difference exists in the production of the different focus types, we predict it will show up in the kind of pitch accent selected by the speakers.

2. METHOD

Three speakers of Florentine Italian (the variety on which Standard Italian is based) read 4 times a set of 4 declarative sentences embedded in brief paragraphs/dialogs. The texts were conceived to convey the appropriate set of presuppositions in order to elicit broad, narrow-semantic and narrow-contrastive focus readings on the same sentence. In each of them the target word associated with the nuclear accent is sentence-final, and varies from (“va”, “Gio’vanni”, “vandali”, “con’validano”). The stressed syllable can be word-initial or word-medial, it can be followed by 0-3 unstressed syllables and it is preceded by an inter-stress interval of 2 or 4 unstressed syllables. The structure of the nuclear syllable is: C_1V or C_1VC_2 , where $C_1=/v/$, $V=/a/$ and $C_2=/n, l/$. The 144 utterances obtained were digitized at 44 KHz, downsampled at 11 KHz and their F0 was extracted using Praat. Targets H and L are defined as follows:

H = the local maximum (peak) at the end of a pitch rise. In a flat plateau, it is defined as the last point; in a slightly falling plateau, as the point at which a rapid fall begins.

L = the local minimum at the end of a fall that keeps constant for at least 3 frames (see fig. 1).

The alignment of H and L is represented as their distance in ms from syllable onset (H/L-SON), from syllable offset (H/L-SOF) and from vowel onset (H/L-VON), and it is expressed as a percentage of syllable or vowel duration.

3. RESULTS

3.1 Broad semantic focus

The present data confirm the results of our previous study.

In 81% of the cases the speakers consistently associate a fall with the nuclear syllable of broad focus declaratives¹. Figure 1 shows how the H and L targets align with the segmental string in the word “con[‘va]lidano”: H consistently aligns with the prenuclear syllable “con”, L with the nuclear syllable “va”.

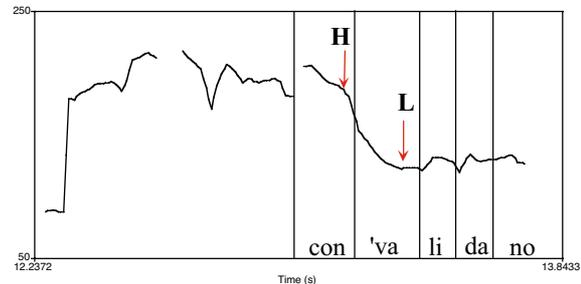


Fig. 1: F0 curve of the broad focus declarative *Hanno detto che lo convalidano* (“They said they validate it”).

Figure 2 shows L-SON and H-SOF for all contexts considered. Independently of the number of postnuclear syllables L is clearly anchored within the nuclear syllable. It aligns no earlier than 58,5% of its duration and no later than 70%. H aligns in the prenuclear syllable at a point in time from syllable offset that corresponds to no less than 16% and no more than 37% of its duration (i.e., H aligns no later than 84% and no earlier than 63% from syllable onset).

The alignment of L and H has been tested with separate one-way ANOVAs performed on pooled data for all speakers. As for L, the dependent variables are *L-SON* and *L-VON*, the independent variable is *number of postnuclear syllables* (4 levels). Results show no significant context effects in the alignment of L either relative to syllable onset ($F(3,30)=1.421$ $p=.256$) or relative to vowel onset ($F(3,30)=2.608$ $p=.069$). However, since *L-VON* almost approached significance at 5%, a series of post-hoc comparisons have been performed. The Fisher’s PLSD post-hoc test (the most liberal) showed a significant difference in L alignment only for the pair “va” vs. “vandali” (0 and 2 unstressed syllables), suggesting a “special” status of the stressed monosyllable in sentence-final position. As for H, we tested how *H-SON* (the dependent variable) is affected by the *position of the prenuclear syllable* (the independent variable) relative to the word boundary. This factor has 2 levels: prenuclear syllable in the target word vs. prenuclear syllable in the preceding word. The results show that H aligns significantly later in the prenuclear syllable when a word boundary follows: $F(1,32)=6,797$ $p=.013$. Moreover, within the subset of words in which prenuclear and nuclear syllables belong to the target word, the timing of

¹ In the remaining 19% of the cases the nuclear syllable is marked by a H target phonologically analyzed as a monotonal peak accent H*. In half of the cases H* is associated with the sentence final monosyllable; in the other half its use is confined to one speaker’s rendition of the sentence “Quei ragazzi sono dei vandali” (“Those boys are vandals”).

H shows no effect of the preceding *inter-stress interval* (two levels: 2-4 unstressed syllables): $F(1,19)=2.98$ $p=.1$.

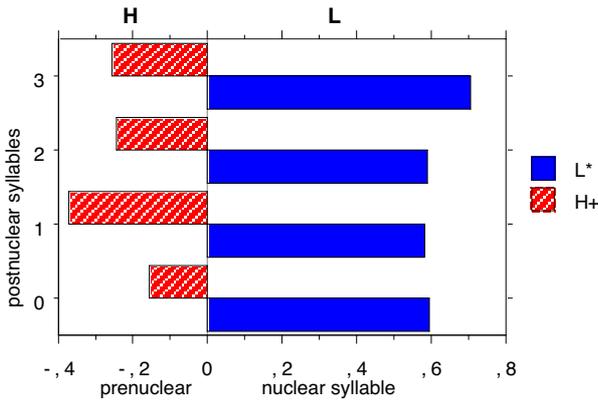


Figure 2. Alignment of H and L in BF declaratives.

3.2 Narrow semantic focus

In 69% of the cases our speakers mark the final accented word with a fall, exactly as in the BF declaratives. A remaining 28% of cases are marked by a peak accent, and a 3% by a steep rise as in NCF declaratives.

3.3 Narrow contrastive focus

Speakers differ in their productions of narrow contrastive focus declaratives. In 15 out of 16 cases one speaker (R) marks both the pre-nuclear and the nuclear syllable with a H target (a peak or a plateau); the other 2 speakers mark the nuclear syllable either with a steep rise that begins just before its onset (speaker S, 16/16 cases) or with a steep rise that starts and ends within the syllable (speaker M, 11/16 cases).

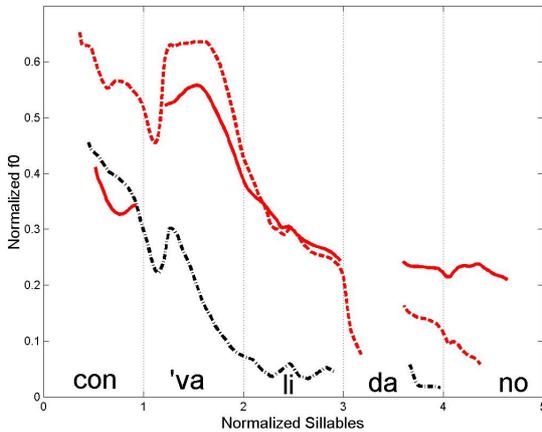


Fig.3: Narrow contrastive focus (speaker R: dashed line; speaker S: solid line) and broad focus (speaker R: dashed-dotted line).

Figure 3 shows the average of 4 F0 curves on the contrastively focused constituent “con[va]lidano” uttered by speaker R (dashed line), and by speaker S (solid line), normalized to syllable duration and pitch range, compared to the nuclearly accented word in the BF sentence by speaker R. For speaker R, separate one-way ANOVAs performed on the alignment of each H target relative to SON showed that: i) the alignment of the nuclear H in the syllable is not affected by the number of post-nuclear

syllables ($F(3,11) = 2,511$ $p = .112$). It aligns no earlier than 31% and no later than 50% of the syllable duration. ii) The pre-nuclear H does not show any effect due to the presence/absence of a following word boundary ($F(3,11) = 0,045$ $p = .834$), nor to the preceding interstress interval: $F(1,13)=0,601$ $p=.452$). Finally, a two-way ANOVA with *tone* (pre-nuclear vs. nuclear) and *number of postnuclear syllables* as factors shows that there is no difference in pitch height between the nuclear and pre-nuclear H: ($F(1,22)= 0,499$ $p=.487$), while there is a significant effect due to the proximity of the edge of the contour ($F(3,22)= 18,835$ $p<.0001$), and a significant interaction ($F(3,22)= 3,563$ $p=.03$). All pairways comparisons are significant except that between words with 2 and 3 following unstressed syllables (Games/Howell post-hoc test).

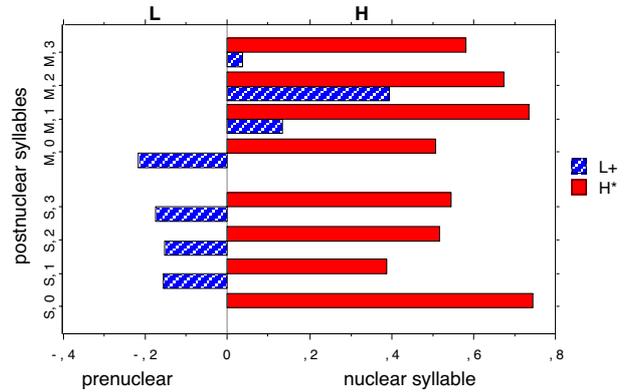


Fig. 4: Alignment of L and H in narrow contrastive focus declaratives.

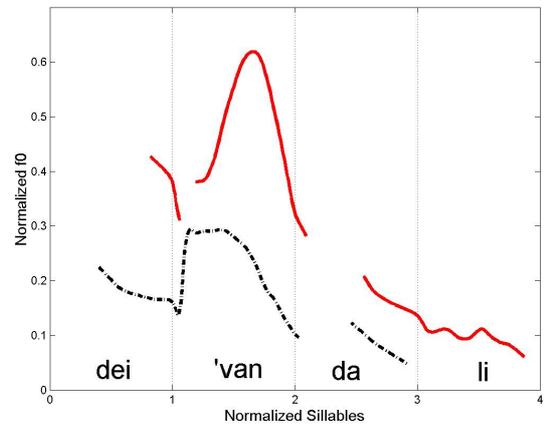


Fig. 5: Narrow contrastive focus (solid line) and broad focus (dashed-dotted line) for speaker S.

For the productions of the other two speakers we measured the alignment of L target with the pre-nuclear and of H target with the nuclear syllable relative to SON. Figure 4 shows the results for speaker M (upper part) and S (lower part). As for the nuclear syllable, both speakers align H within the syllable (no earlier than 46% and no later than 67% of syllable duration). A two-way ANOVA with *number of postnuclear syllables* and *speaker* as factors shows no effects of the main factors (respectively: $F(3,19)= 0.583$ $p=0.633$; $F(1,19)= 3.302$ $p=0.085$) and a significant interaction ($F(3,19)=7.519$ $p=0.016$). However, the speakers differ in their alignment of L. L always aligns

with the prenuclear syllable for speaker S (no earlier than 82% and no later than 100%); L aligns or with the prenuclear (in the monosyllable) or with the nuclear (in the polysyllables) for speaker M.

4. DISCUSSION

The alignment data confirm our previous analysis of the nuclear fall in BF declaratives as a H+L* pitch accent. The accentual nature of L is supported by its consistent anchoring within the stressed syllable. If the L target were a boundary and not an accentual tone, we would expect the number of following unstressed syllables to affect its timing, progressively moving its anchoring far from the stressed syllable and towards the end of the word. This prediction was not borne out.

NCF declaratives are realized with a sustained high or as a rising nuclear accent. One speaker (R, 31% of the total cases) produces a peak (or plateau) accent preceded by a plateau on the prenuclear syllable; the other two speakers (S and M, 56% of the total) produce a peak accent preceded by a marked valley. The details of alignment and scaling for the two H targets for speaker R yield to a representation of a bitonal H+H* accent. Even if this representation obviously violates OCP, pitch accents with tones of the same polarity have been already proposed in the literature [11] (and see [1] for other references).

The steep rise produced by S and M is here analyzed as L+H* and not as H* as in [9]. The difference between H* and L+H* is one of the most problematic to assess, and has received a considerable amount of attention [1, 12]. A comparison between rising accents produced on the same word by the same speaker S in different focus structures supported our decision. Fig. 5 shows the normalized F0 curves for the final phrase of BF² (dashed-dotted line) and of NCF (solid line) renditions of the sentence "Quei ragazzi sono dei vandali". It can be seen that the pitch peak in the NCF accent is higher than the end of the plateau in the BF accent. But also that the alignment of the peak in NCF is later in the syllable. It could be thought that the later alignment is due to the longer time it takes to complete a higher pitch rise. Indeed, a pitch rise of 3,1 semitones in BF is accomplished in 102 ms, while a pitch rise of 5,47 semitones in NCF is accomplished in 170 ms. However, if we apply to these data the equation proposed in [13] to calculate the maximum speed of pitch change, we obtain that the time taken to complete the rise in NCF is longer than the minimum time needed to complete a rise of 5,47 semitones: 137 ms vs. 170. Hence, we can rule out the possibility that the later alignment is due to an articulatory constraint and capture the difference between the two accents describing the rise in NCF as L+H* and that in BF as H*.

Finally, the focused constituent in NSF declaratives is realized with the same H+L* pitch accent used in BF declaratives in the vast majority of the cases (69%).

5. CONCLUSION

Our study showed that in Florentine Italian semantic focus and contrastive focus are signaled by different pitch accents. We have seen that in the vast majority of the cases broad F is signaled by a bitonal pitch accent with a L target, H+L*, confirming [9] and [5]. Contrastive F is signaled by a bitonal pitch accent with a H target: either a rising L+H* accent or a sustained high H+H* accent. This result is compatible with findings for English [2,3] while it is not compatible with findings for Dutch [4]. On the other hand, broad and narrow semantic F are not distinguished via pitch accent type: in both cases the last focused constituent is marked by the same H+L* pitch accent in the vast majority of cases. This can lead to ambiguity in focus scope.

REFERENCES

- [1] Ladd, R., *Intonational Phonology*. Cambridge: Cambridge University Press, 1996.
- [2] Pierrehumbert J. and Hirschberg J., "The meaning of intonation contours in the interpretation of discourse". In Cohen P., Morgan J. e Pollack M. (eds.), *Intentions in Communication*, Cambridge MA, MIT Press, 1990.
- [3] Selkirk E., "Contrastive FOCUS vs. presentational focus: Prosodic evidence from right node raising in English", *Speech prosody 2002*, pp 643-646.
- [4] Khramer E. and Swerts M, "On the alleged existence of contrastive accents", *Speech Communication*, 34, pp. 391-405, 2001.
- [5] Grice, M., D'Imperio M, Savino, M. and Avesani C. Towards a strategy for ToBI labelling varieties of Italian. In S-A. Jun (Ed.), *Prosodic Typology and Transcription: A Unified Approach*, in press.
- [6] D'Imperio, M. and House D., "Perception of questions and statements in Neapolitan Italian. In G. Kokkinakis, N. Fakotakis, and E. Dermatas (Eds.), *Proceedings of Eurospeech'97*, Vol. 1, pp. 251-254, 1997.
- [7] Frota, S., *Prosody and Focus in European Portuguese*, New York: Garland, 2000.
- [8] Swerts M, Krahmer E. and Avesani C., "Prosodic marking of information status in Dutch and Italian: a comparative analysis", *Journal of Phonetics*, 30, pp. 629-654, 2002.
- [9] Avesani C. and Vayra M., "Costruzioni marcate e non marcate in italiano. Il ruolo dell'intonazione" *Atti delle X Giornate di Studio del GFS*, Napoli, 13-15 dicembre 1999, 2000, 1-14.
- [10] Nespor M. and Guasti M.T., "L'allineamento tra focus e accento", *Lingue e Linguaggi*, 1, 2002.
- [11] Grice M., "Leading tone and downstep in English", *Phonology*, 12, pp.183-233, 1995.
- [12] Ladd R. and Schepman A., "'Sagging transitions" between high pitch accents in English: experimental evidence", *Journal of Phonetics*, 31, pp. 81-112, 2003.
- [13] Xu Y., "Articulatory constraints and tonal alignment", *Speech prosody 2002*, pp 91-100.

² See footnote 1