

Accenting deaccenting and information structure in Italian dialogue

Cinzia Avesani

ISTC-CNR — Italy

(Institute for Cognitive Sciences and Technologies)

avesani@csrf.pd.cnr.it

Mario Vayra

University of Siena - Italy

vayra@unisi.it

Abstract

Do Italian speakers deaccent given information? In this study we examined word tokens repeatedly mentioning the same entity within a task-oriented dialogue. Contrary to what is expected in Germanic languages, results show that the vast majority of the repeated mentions are accented irrespective of their being hearer/discourse given or discourse segment given.

Keywords: information structure; prosody.

1 Introduction

Deaccenting is defined as the absence of a pitch accent on a word that might otherwise be expected to be accented (Ladd, 1980; Swerts, Krahmer and Avesani, 2002). While accenting is normally used by a speaker as a prosodic pointer to information that he/she intends to present as new to the listener or as a pointer to a contrast-relation, deaccenting may be used to signal that a word refers to given information (or information that can otherwise be expected in the discourse). This commonly held view rests on plenty of linguistic and psycholinguistic evidences for a close correspondence between information status and (de-)accentuation in Germanic languages (Nootboom and Terken, 1982; Terken, 1984; Fowler and Housum, 1987; Nootboom and Kruijff, 1987; Terken and Nootboom, 1987; Horne, 1990; Hirschberg, 1993). However, few studies have shown that a previous mention in the discourse does not suffice to trigger deaccentuation nor in controlled experimental settings (Terken and Hirschberg, 1994) nor in spontaneous dialogue (Bard and Aylett, 1999).

Stronger evidence against deaccentuation as a general phenomenon comes from Romance languages. In a cross-linguistic study, Cruttenden (1993) reports that certain Romance languages as Catalan and Spanish

resist deaccentuation. Ladd (1996) acknowledges that Italian and Rumanian also resist deaccentuation, and Gussenhoven (2004) classes French among the languages with unfrequent deaccentuation.

Italian patterns with Germanic languages as far as deaccenting of clauses and simple NPs (Avesani, Hirschberg and Prieto, 1995; Hirschberg and Avesani, 1997; D Imperio, 1997; Farnetani, and Zmarich, 1997); but differs from those in avoiding deaccenting within NPs (Swerts, Krahmer and Avesani, 2002). Moreover, differently from English, items that have been previously mentioned in the same discourse segment can be accented again irrespective of grammatical function and surface position in the sentence in samples of broadcast speech (Avesani, 1997).

In this study we examined word tokens repeatedly mentioning the same entity within a task-oriented dialogue. Each co-referential repetition was coded according to its syntactic category, grammatical function, position in the utterance and in the prosodic phrase, focus status and information status.

A possible definition of information status rests upon the notion of shared knowledge of discourse participants (e.g. the taxonomy of shared knowledge in Prince, 1981). An entity may be informationally given or new with respect to (the speaker's beliefs about) the hearer's belief: hearer-new entities are assumed not to be already known to the hearer and are typically in the linguistic form of indefinite referring expressions, while hearer-old entities are typically definite. An entity may be old/new also with respect to the discourse model. Discourse-newness tells us nothing about an entity's hearer status, but a hearer—new entity would tell us something about its discourse-status: it would be necessarily discourse-new, since hearers are expected to remember what they have been told (Prince, 1992).

In a different perspective, we can characterize an entity's information status relatively to a dynamically unfolding record of mutual beliefs established during the discourse (Grosz and Sidner, 1986). According to this theory of discourse structure, discourse is

comprised of Discourse Segments (DS), whose hierarchical relationships are determined by the intentional structure and realized by the linguistic structure. A third structure, the attentional state, is responsible of the saliency of an entity in a DS. The attentional state directly refers to the discourse structure (the intentionally-based hierarchy of DSs) to determine which entities are accessible or salient to the discourse participants at any given point in time: as the discourse and the intentions driving it evolve over time, so does the model of the attentional state, and the entities which are salient within this model also dynamically change over time. The notion of saliency is formalized by attentional state mechanisms that are claimed to underlie discourse processing in general.

Interesting for us is the *global* level of attentional state, which is modeled as a last-in first-out Focus Stack of Focus Spaces, each containing representations of the entities and relations salient within the discourse segment corresponding to the focus space. When a discourse is initiated, a focus space is pushed onto the empty stack, representing the first DS and its purpose. Elements contained within that DS are added to the stack and recorded as they are mentioned until the discourse segment closes and its focus space is popped from the focus stack. When entities are newly added to the focus space they are not globally salient yet; when they are already in the space they are considered salient. A salient entity is activated in the speaker's awareness and is, therefore, given.

2 Method

2.1 The dialogue

The dialogue we choose to analyze is taken from the IPAR corpus of spoken Italian (Albano Leoni and Giordano, forthcoming). Two speakers of the Roman variety of Italian, one male and one female, have to cooperate in order to find the differences in two slightly

different versions of the same picture (figure 1). They have been recorded in a silent booth and cannot see each other. The left picture has been assigned to speaker 2, who initiates the dialogue and has the role of instruction giver; the right picture has been assigned to speaker 1 who acts as follower.

The dialogue consists of 336 turns. The task at hand favors the spontaneous production of word tokens repeatedly mentioning the same entity. We found 36 repeated entities (e.g. the boat, the child, the ball etc.) and 294 co-referring expressions (e.g. the ball, the balls, the black ball etc.), including the absolute first mention of each entity. For each entity the number of repetitions vary from a minimum of 2 to a maximum of 41.

2.2 The analysis

Each co-referring expression has been coded according to a number of phonological, syntactic and pragmatic variables. Those include: word's syllabic structure; syntactic category (POS); grammatical function (subject, direct/indirect object); change of grammatical function from previous mention; focus type; information status (given/new with respect to the hearer and discourse and given/new with respect to the discourse segment); position in the utterance and in the prosodic phrase (final vs non-final); occurrence in echoic responses.

The utterances in which the co-referring expressions occur have been at first perceptually analyzed by an independent phonetician in order to identify the presence or the absence of an accent. Then, the utterances have been ToBI transcribed by the same phonetician and by the first author. Difficult cases that gave rise to different analyses (for example expressions judged by one transcriber as deaccented and by the other as associated with L* or with a highly downstepped H* accent) have been discussed until a consensus decision could be reached.



Figure 1. The dialogue task: to identify the differences in the picture.

For each expression the following prosodic variables have been considered: presence/absence of pitch accent; type and role of pitch accent (nuclear vs. prenuclear); level of nuclearity (nuclear in intermediate or intonational phrase); type of following boundary tone; position in the prosodic domain (final vs. non final).

3 Results

3.1 How much deaccenting?

Out of 294 repeated mention of the same entity, two word tokens were discarded. Of the remaining 292 repeated expressions analyzed, 273 are accented (93%), 19 are deaccented (6.5%). Even if speaker 2 (the instruction giver) deaccents less than speaker 1 (the follower) - respectively 7 vs 12 cases - the difference is not statistically significant ($\chi^2 = 2,969$, $df=1$, $p=0.0849$).

3.2 Deaccenting and information status

We used two operational definitions of the information status of a referring expression. The first, based on the notion of shared knowledge, define as new the absolute first mention of an entity in the entire discourse; given any other mention of it is given. We will refer to this definition as static given.

The second definition crucially refers to the discourse structure of our dialogue. Three independent labelers have segmented the dialogue in DSs and represented their hierarchical relationships. The instructions for annotating the dialogue have followed those presented by Nakatani, Grosz, Ahn and Hirschberg (1995). The structure in figure 2 is derived from their consensus. We define as new the first mention of an entity in a DS; any other mention in the same DS is treated as given. If an entity has an antecedent in a previous DS, it is new if its antecedent is in a different DS of the same level of embedding (e.g. current mention in DS7, antecedent in DS6); it is given if its antecedent occurs in hierarchically dominating DS (e.g. current mention in DS5, antecedent in DS4). We will refer to this definition as dynamic given.

The data of both speakers were used to verify if a significant relation exists between the accenting of a word and its being statically or dynamically given. Results show that no significant relation emerges between accentual status and information status, nor when an entity is considered given with respect to the hearer and to the discourse (statically given: $\chi^2 = 0.001$, $df=1$, $p=0.97$) nor when it is considered as

given with respect to its belonging to a DS (dynamically given: $\chi^2 = 8.107$, $df=1$, $p=0.087$).

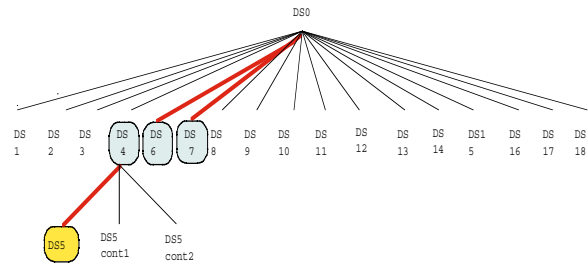


Figure 2. The dialogue discourse structure based on the consensus labelling of three independent labelers.

3.3 Deaccenting and grammatical function

Following Terken e Hirschberg (1994), we asked whether the massive accenting of given information is related to the fact that a repeated entity changes its grammatical function from the immediately previous mention. Both speakers tend to refer to an entity maintaining the same grammatical function of the previous item (speaker 1: 49 vs 94 cases; speaker 2: 36 vs 85 cases), but changing the grammatical function does not significantly relates to the accentual status of the repeated item ((speaker 1: $\chi^2 = 1.427$, $df=1$, $p=0.232$; speaker 2: $\chi^2 = 0.861$, $df=1$, $p=0.353$).

3.4 Information status and prosodic form

Even if speakers do accent given information, they could still distinguish givenness and newness by using different types of pitch accents (Hirshberg and Pierrehumbert, 1990). In Italian different types of pitch accents distinguish items in presentational and contrastive focus in sentence final position (Avesani and Vayra, 2004). In the same vein, it would be perfectly possible that a particular pitch accent is systematically devoted to mark given information. As a matter of fact, in a sample of spontaneous (but quite formal) broadcast speech, Avesani (1997) did find that L* accents were associated only with informationally given items and, in samples of read speech, with items that are formally external to the proposition, such as vocatives (Avesani, 1995).

However, the present analysis shows no significant relation between the type of pitch accent used on accented items and their information status. The pitch accents attested in the dialogue do not distribute differently on items that are statically or dynamically given from items that are new (respectively: $\chi^2 = 3.457$, $df=4$, $p=0.484$; $\chi^2 = 8.107$, $df=4$, $p=0.087$). The frequency of the different pitch accents on dynamic given items is shown in figure 2.

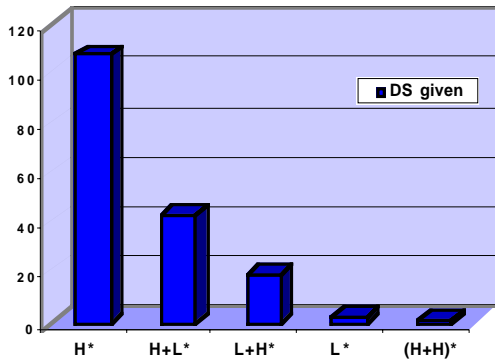


Figure 2. Distribution of pitch accent types on dynamically given items.

Finally, we investigated whether our speakers used a difference in the duration of the accented items to distinguish their information status. A two-way ANOVA with factors Word Structure (bisyllables vs trisyllables) and Information Status (static-given vs static new) showed that there is a difference and that it is significant: $F(1,279)=5.853$, $p=0.016$. However, a separate ANOVA calculated for each speaker revealed that the difference is entirely due to the contribution of speaker 2 ($F(1,123)=7.42$, $p=0.007$), while for speaker 1 the difference is non significant ($F(1,152)=1.59$, $p=0.207$). The data are shown in figure 3.

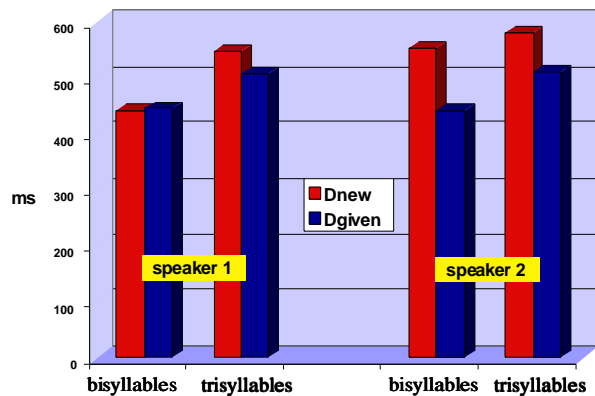


Figure 3. Duration of the accented words coded as statically given/new.

Moreover, when the factor Information Status refers to given/new in a DS, the result of the analysis of variance applied to the data of both speakers turns to non significant: $F(1,279)=1.594$, $p=0.207$.

The only significant difference in duration of an accented or deaccented word is related to its position in the prosodic domain ($F(3,275)=36.96$, $p<0.0001$). Figure 4 shows that final lengthening applies to an

item flanking the right boundary of a prosodic domain independently of its accentual status.

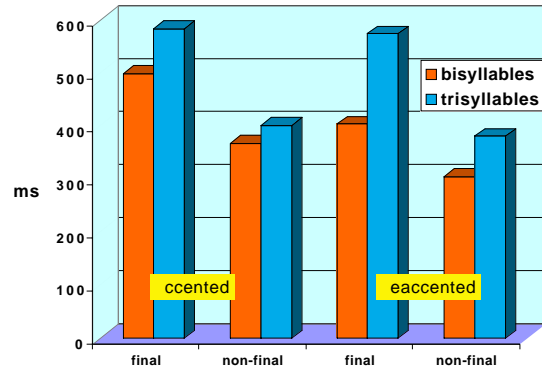


Figure 4. Duration of the accented words according to their position in the prosodic phrase.

4. Discussion and conclusion

This analysis of a task-oriented spontaneous dialogue shows that in Italian deaccenting is very rare. More rare than in the comparable English dialogues analyzed by Bard and Aylett (1999): 6.5% in Italian vs. 20-25% in English).

Deaccenting is not related to the informational status of an entity. Speakers do assign a pitch accent to co-referential expressions that are informationally given, whichever is our definition of givenness: the shared knowledge of the conversation participants (Prince 1981, 1992) or to the dynamically unfolding records of mutual beliefs established during the discourse (Grosz and Sidner, 1986). Our speakers do not mark the repeated mentions of the same entity as given to the hearer nor by using a special type of pitch accent, as shown for English by Pierrehumbert and Hirschberg (1990), and as occasionally appeared in our previous data of more formal speech (Avesani, 1997), nor by lengthening their duration as shown by Fowler and Housum (1987).

Still, deaccenting does occur on a minority of repeated expressions, and it is worth exploring their linguistic nature. Deaccented items falls in four categories:

i) Postfocal deaccenting.

9 out 19 cases belong to this category. Deaccenting occurs after a) contrastive focus, b) counterpresupposition focus (Gussenhoven, forthcoming) and c) presentational focus. In the following excerpt, we give an example of contrastive and counterpresupposition focus deaccenting.

contrastive focus. Context: speaker 2 drives speaker s 1 attention to the ball, specifically to the

signs connecting the three black spots on the ball.

Sp1: con tre segnetti?

with three little signs?

Sp2: s / con [DUE]_{Fcontr} L+H* **segnetti** (deacc)

yes, with [TWO]_{Fcontr} (L+H)* **signs-DIM** (deacc)

counterpresupposition focus. Same context.

Sp1: no [IO L+H* non PENSO

L+H*]_{FOC} che questo sia attaccato al

bordo, il terzo **segnetto** (deacc)

No, I H* do not **think** L+H* that this is attached to the border, the third **sign-DIM** (deacc)

ii) *deaccenting of the first item in a informationally omogeneous complex NP:*

Context: speaker 2 initiates the dialogue with the description of a sailing boat. After few turns she asks:

Sp1: Giordano, tu come ce / tu come lo vedi il disegno, scusami.

Giordano, in which way do you look at the the picture, excuse me

Sp2: io c'ho no, io c'ho io<oo> / come la vedo io/ c'ho a destra questa [barca (deacc) a vela]_{NP, given} H+L*

*I have, no, I have/ how do I see it/ I have on my right this boat (deacc) for sailing H+L**

iii) *deaccenting of subjects as topic (one case only):*

Sp2: a che altezza sta la nuvola<aa> rispetto alla<aa> alla<aa> <mh> alla barca?

at which height is the cloud with respect to the the the boat?

Sp1: quale nuvola?

Which cloud?

Sp2: [la **nuvola** (deacc)]_{TOPIC} [quella di **destra** H+L*]_{FOCUS}

*the cloud (deacc) the one on the right H+L**

iv) *deaccenting in echoic response.*

Sp2: <mh> sotto, [i capelli], ci sono sotto l'orecchio?

below, the hair, is there any hair below the ear?

Sp1: s, [po' di capelli], s, capelli neri s

yes, a little hair (deacc), yes, black hair yes

Echoing is defined as a speaker's lexical repeat of (parts of) an utterance spoken by a conversation partner in a previous turn. Strictly speaking, repeated lexical items in these contexts cannot be said to be informationally given: they carry significant information related to the conversation process itself and are significant meta-linguistically; the information they carry cannot be said to be related to what the speaker explicitly intends or to what the listener consciously utilizes (Shimojima, Katagiri,

Koiso and Swerts, 2002). All other cases of deaccenting are structure-driven: they appear to be constrained by the position of the current item in a focus domain or in a syntactic phrase.

Our data contrast the view of deaccenting as a cognitively-based universal phenomenon and support Ladd's intuition of a distinction between languages that permit, prefer or require the deaccenting of repeated lexical material, as Germanic languages, and languages in which such deaccenting is dispreferred or syntactic-constrained, primarily achievable through word order modifications as Romance languages.

References

- Albano Leoni F. and Giordano R. (eds.) (forthcoming). *Italiano parlato. Analisi di un dialogo*, Napoli, Liguori (Quaderni di Comunicazione parlata, 1).
- Avesani, C. (1995). ToBI. Un sistema di trascrizione per l'intonazione italiana, *Atti delle V giornate di Studio del Gruppo di Fonetica Sperimentale*, Trento, novembre 1994, pp. 85-98.
- Avesani, C. (1997). I toni della RAI. Un esercizio di lettura intonativa. In *Gli italiani trasmessi: la radio*. Firenze: Accademia della Crusca, pp. 659—727.
- Avesani, C., Hirschberg, J. and Prieto, P. (1995) The intonational disambiguation of potentially ambiguous utterances in English, Italian and Spanish. In *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, pp. 174—177.
- Avesani C. and Vayra M. (2004). Focus ristretto e focus contrastivo in italiano. In Albano Leoni F., Cutugno F., Pettorino M, Savy R: (a cura di), *Atti del Convegno Il parlato italiano*, Napoli, 13-15 febbraio 2003, M. D Auria Editore, CD- ROM, F01, pp.1-20.
- Bard, E.G. and Aylett, M.P. (1999) The dissociation of deaccenting, givenness, and syntactic role in spontaneous speech. In *Proceedings of the XIVth International Congress of Phonetic Sciences, San Francisco*, pp. 1753—1756.
- Cruttenden, A. (1993). The de-accenting and re-accenting of repeated lexical items. In *Proceedings of the ESCA workshop on Prosody, Lund*, pp. 16—19.
- D Imperio, M. (1997) Narrow focus and focal accent in the Neapolitan variety of Italian. In *Proceedings of the ESCA workshop on intonation*, Athens, pp. 87—90.

- Farnetani, E. and Zmarich, C. (1997) Prominence patterns in Italian: an analysis of F0 and duration. In *Proceedings of the ESCA workshop on intonation*, Athens, pp. 115—118.
- Fowler C. and Housum J. (1987). Talkers signalling of new and old words in speech and listeners perception and use of the distinction, *Journal of Memory and Language*, 26, pp. 489-504.
- Grosz B. and Sidner C. (1986). Attentions, intentions and the structure of discourse, *Computational Linguistics*, 12, pp. 175-204.
- Gussenhoven C. (2004). *The phonology of Tone and Intonation*, Cambridge, Cambridge University Press.
- Gussenhoven C. (forthcoming). Types of focus in English. In B ring, Gordon and Lee (eds), *Topic and Focus: papers from a workshop on Intonation and Meaning*, Dordrecht, Kluwer.
- Pierrehumbert J. and Hirschberg J. (1990). "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.
- Hirschberg, J. and Avesani, C. (1997) The role of prosody in disambiguating potentially ambiguous utterances in English and Italian. In *Proceedings of the ESCA workshop on intonation*, Athens, 1997, 189—192.
- Horne M. (1990). Accentual patterning in New vs Given subjects in English, Lund University, Dept. of Linguistics, Working Papers 36, pp.81-97,
- Ladd, R. (1980) *The structure of intonational meaning*, Bloomington, Indiana University Press.
- Ladd R. (1996). *Intonational phonology*, Cambridge, Cambridge University Press.
- Nakatani C., Grosz B., Ahn D. and Hirschberg J. (1995). Instructions for annotating discourses, Technical report TR-21-95, Center for Research in Computing Technology, Harvard University.
- Nooteboom S.G. and Terken J:M:B: (1982). What makes speakers omit pitch accents? An experiment, *Phonetica*, 39 (4-5), pp. 317-336
- Nooteboom S.G. and Kruyt J.G. (1987). Accent, focus distribution, and the perceived distribution of given and new information: an experiment, *Journal of the Acoustical Society of America*, 82, pp. 1512-1524.
- Prince E.F. (1981). "Toward a taxonomy of given-new information". In Cole P. (ed), *Radical pragmatics*, New York, Academic Press, pp. 223-255.
- Prince E.F. (1992). The ZPG letter: subject, definiteness and information status. In Thompson S. e Mann W. (eds), *Discourse description*, Amsterdam, Elsevier Science Publishers, 295-325.
- Shimojima A., Katagiri Y., Koiso H. and Swerts M. (2002). Informational and dialogue-coordinating functions of prosodic features of Japanese echoic responses, *Speech Communication*, 36:1/2, 113-132.
- Swerts M., Krahmer E. and Avesani C (2002). Prosodic marking of information status in Dutch and Italian: a comparative analysis , *Journal of Phonetics*, 30, 4:629-654.
- Terken, J. (1984). The Distribution of Accents in Instructions as a Function of Discourse Structure, *Language and Speech*, 27, 3, 269-289.
- Terken, J. and Nooteboom, S.G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for Given and New information, *Language and Cognitive Processes* 2, 145-163.
- Terken, J. and Hirschberg, J. (1994) Deaccentuation of words representing given information: effects of persistence of grammatical function and surface position, *Language and Speech*, 37(2), 125—145.