

Perlocutory Values and Functions of Melisms in Spontaneous Dialogue

Geneviève Caelen-Haumont

Laboratoire Parole et Langage, UMR CNRS 6057
Université de Provence, Aix-en-Provence, France
gcaelen@lpl.univ-aix.fr

Abstract

Following earlier work [3][5][6], this study, based on a corpus of spontaneous dialogue, aims to define melisms as opposed to intonation and to specify not only their forms, but also their meanings and functions. Melisms appear to be directly related to the domain of beliefs, of subjective values and hence of their perlocutory effects. The study shows in particular how melisms, beliefs and motivations are organised in the discourse of speakers.

1. Introduction

In the last 50 years a great number of studies have been carried out in the field of prosody. In the early period, as is well known, the emphasis was placed on the influence of syntax, particularly in the study of read speech. In the '80s the importance of semantics and of pragmatics was recognised more and more, particularly in the study of spontaneous dialogue. Studies of the importance of emotions have also a similarly long history [8] [9], and today are attracting more and more attention. The work of Cosnier [7] for example, is the result of an interaction between theory of enunciation and emotion. The majority of studies, however, have been concerned with the study of so-called "secondary" emotions [11], without linking these to everyday speech. This paper aims at shedding light on what we call "ordinary emotion" and its prosodic and melodic correlates, *i.e.* melisms and their values and functions.

2. Intonation and melism

In our perspective, intonation and melisms are two distinct prosodic forms. The notion of "melism" is borrowed from the domain of singing and refers to a melodic figure spread over the duration of the word, such that the number of notes perceived is higher than the number of syllables in the word. Applied to the domain of language, this concept refers to the pitch excursions specific to speech (as well of course as to song).

2.1. Prosodic Forms

A melodic focus is an example of a case of melism. We do not use the term focalisation for two reasons. First of all, the term is inadequate since uses a binary concept (focal/non-focal) for what is in reality a scalar process [4][13]. Secondly, the term is vague and confuses the domains of acoustics, semantics and pragmatics. In reference to prosodic form we consequently prefer to use the term melism. When speaking of prosodic content, we will refer to semantic and pragmatic value. In the case of the beliefs attached to these values and their effect on the listener, [1], we use the expression "perlocutory aims".

On the level of the pitch contour of the melism, the common factor of the different forms observed is the use of the High (or even the extra-High) pitch level^{*}. Two types of melism are observed, one where all the pitch modulations are restricted to the High level (in which case the pitch contour can sometimes take the form of a plateau), and another where the pitch modulations take the form of a large movement from the Low level to the High or Extra-High level. The significant pitch movement generally takes place within the range of the middle two levels, (Mid-High and Mid-Low) across a dozen or more semi-tones. Apart from the plateau produced in the high level (see below, fig. 3 "prenez"), four different forms of melism are found within the word: a simple slope, ascending or descending (fig. 2, "très" "très"), parallel slopes (fig.1, "patiné"), alternating slopes which can be circumflex, straight (fig. 1, "récent") or reversed (fig. 3, "souris"), and finally mixed cases. In our examples melisms are produced across a range of from 8 to 12 semitones, in a register ranging from 180 Hz to a maximum of 410 Hz.

2.2. Melisms and intonation

In their relation to intonation, melisms play two roles. Situated at the right edge of a group, a melism combines with the intonation and contributes by a significant pitch movement and a slowing down of speech rate to the prominence of the syntactic or syntagmatic boundary and/or to the semantic or pragmatic value of the final lexical item of the group. Situated at an initial or an internal position in the group, the pitch movement, accompanied or not by a pause, lends a equally large prominence to the lexical word. This generally creates a break, a discontinuity within the pitch pattern associated with the word (parallel pitch movements for example) and/or with respect to the context. This process is then often manifested by 'disyntactic' effects (such as tonal disjuncture and pause between two syntactically dependent units) and 'asyntactic' effects such as the grouping together of two units which are syntactically independant. [2],[5]. It should be noted that melisms may include patterns corresponding to emphatic accents but may also be associated with non-emphatic patterns.

2.3. Associative vs. dissociative forms

Prosody in our perspective is characterised by the interaction of two prosodic processes, one the support of linguistic structure, the other the breakdown of this support; the first converging towards the group, the second towards the elementary, the local. This mechanism, in fact, gives rise to

* We assume here following the tradition since [13] that pitch can be described on the basis of five speaker dependent pitch levels or registers which we refer to (from highest to lowest) as Extra-High (*supra-aigu*), High (*aigu*), Mid-High (*infra-aigu*), Mid-Low (*supra-grave*) and Low (*grave*).

two types of contradictory forces: intonation actualises the forces of association, while lexical prosody actualises the forces of dissociation. In the domain of physics, this process is translated in terms of entropy of the system when it evolves from order to disorder, giving more liberty, initiative, choice, hesitation (melism), and its opposite, or negentropy, when the system evolves towards a degree of increasing organisation (intonation).

2.4. “Ordinary emotion”

The literature traditionally contrasts two main categories of emotions [11], on the one hand “crude emotion”, or “primary emotion”, which is of a physiological nature and on the other hand “socialised emotion” or “secondary emotion”, mastered in language. Here we make room for another type of emotion, “ordinary emotion”, which is in fact the mechanism at the root of the subjective being and of its overall sensations and expressions. This ordinary emotion linked either to profound or to more superficial beliefs and values of the individual, is always at work in discourse because of this very fact.

While this ordinary emotion places specific marks in the signal, and therefore in the prosodic contours, it has the characteristic of not leaving any trace of an emotion typified as joy, surprise, anger etc, since these refer to a primary or secondary emotion. Ordinary emotion is simply the expression of attachment to a feeling, an idea, or an intention towards a person, a being, an object, or an idea, this attachment being expressed by a word characterising or evoking it to the best in the utterance, according to the speaker’s motivations or values.

2.5. The subjective space of prosody

In our perspective, the subjective space of the prosody of speech is defined with the help of a certain number of criteria [5]:

- *pitch cues*: maximum of F0 and melisms,
- *a dimension*: the lexical word, more rarely the grammatical word or morpheme,
- *a pitch structure* (and associated parameters): it relies on a scale of a subjective nature, the F0 amplitude, the greatest attributed to a word conveying a very strong subjective value in the utterance,
- *a metalanguage*. This metalanguage, etched in the space of the word, implicit by nature, interposes a local subjective filter vis a vis the linguistic structure, which, in a manner more or less marked in the utterance, situates the speaker with his beliefs, motives, impulses, propensities and dislikes. It is a point of view from within that expresses itself, and in relief, in the foreground. It is a prosodic metacontent of subjective expression with metalinguistic, pragmatic and/or psychological motivations,
- *a function*, that of appropriation of the act of speech.

3. Corpus

12 speakers (6 men, 6 women) recorded, in pairs, 6 spontaneous dialogues of 15 mn each in a sound-proof room. The speakers were placed back to back. The digital recording was carried out on separate tracks for each speaker.

The dialogue produced was a role play between two persons (a tourist and an employee of the local tourist office)

which the speakers produced after a preliminary practice session. Each speaker had a town map in front of them: the maps were partially different from each other, one containing new names for some streets or buildings. On the one hand, the dialogue aimed at the resolution of specific tasks, the speakers were instructed: 1/ to update the town map 2/ to establish a program for a visit 3/ to fix the itinerary for these activities. The dialogue also involved the management of a conflict of objectives since while the tourist was told to express interest in sporting activities the employee was told to insist on cultural visits. An itinerary had to be drawn up in accordance with the town as it is today taking into account road-works, modified one-way streets, etc.

The present study bears on the output of a single speaker’s, in this case a female speaker (O4).

From the constraints bearing on the dialogue as defined above, we propose the following hypotheses:

- (i) the speaker is subjectively involved in her role of employee of the tourist office,
- (ii) this involvement will be prosodically marked by melisms,
- (iii) melisms will affect the lexical items which are related to the objectives of the dialogue and the personal objectives of the speaker.

The objectives of the dialogue can be identified through a number of semantic and pragmatic lexical fields which follow from the above tasks:

- A) the spatial coordinates of the town and the tourist objects (names and directions),
- B) the elements of spatial, temporal and economic description of the places and activities (tourist objects and their characteristics),
- C) the task of updating the map, the computer tools and their manipulation (technical task, map symbols),
- D) evaluations linked to criteria of choice of places and activities (subjective qualities attributed to the objects) or to the technical (appreciative) task of the tourist.

It should be noted that points B and D contribute both to the objectives of the dialogue and to the convergent objectives of the speaker.

4. Analysis of the melisms

For this study we identified by means of an auditory analysis the complete set of melisms produced by speaker O4.

Out of a total of 1860 words (including both grammatical and lexical words) produced by the speaker, 343 lexical words were pronounced with a melism. Among these, we made a distinction between the lexical items at the right edge of a group and the others. Melisms occurring at the end of a group do not isolate the subjective process since this can be combined or not with the syntactic process of lending prominence to the terminal boundary.

Because of this, in this study we shall be particularly concerned with melisms occurring within a group. For these melisms, the demarcative function and the syntactic relations were subordinated to the information function and it is easier to characterise their function and correlates.

It is in these cases that the mechanism of subjective appropriation appears most clearly. The melisms introduce melodic and prosodic interruptions into the syntactic thread. In the corpus produced by speaker O4, 56% of the melisms corresponded to a word (a lexical word or an enunciative particle) in group-final position and 44% to a group-internal lexical word.

4.1. Melisms and lexical fields

In table 1 below, categories A, B, C and D refer to the points described in paragraph 3 above. Category E refers to vocabulary of discursive exchange (verbs and adverbs of interaction, phatic particles) and category F refers to other elements (prepositions with melisms of hesitation).

Table 1: Percentage of melisms in function of the semantic/pragmatic fields of the lexical words

| A | B | C | D | E | F |
|-----|-----|----|-----|-----|----|
| 17% | 35% | 2% | 17% | 27% | 2% |

It can be seen immediately from this table that in this corpus, a large majority of the melisms (71%) are linked to the semantic/pragmatic fields expressing the objectives of the dialogue. From this point of view we may consequently conclude that the hypotheses proposed in paragraph 3 above were verified. The remaining melisms (29%) were mostly linked to the lexicon of discursive exchange.

As we can see the melisms correspond:

(i) to the elements of description of objective tourist objects (B, e.g. “église”, “tourelle”, “menu”...). Figure 1 below presents two examples of these elements with melisms on the words “récent” (hat-shaped about twelve semi-tones) and “patiné” (parallel slopes about 8 semi-tones).

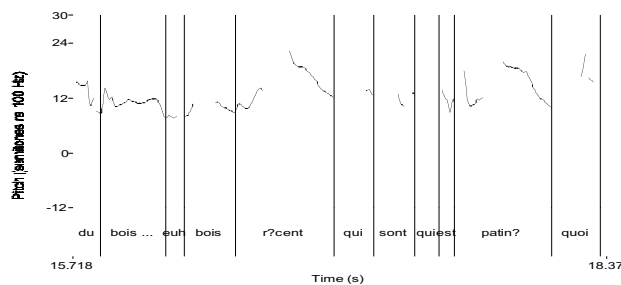


Figure 1: melisms on the words “récent” and “patiné”

(ii) to elements of subjective appreciation (D, e.g. “traditionnel”, “joli”, “dépayçant”, “c’est bien”...), these apply either to tourist objects or to the respective tasks as perceived by speaker O4. Figure 2 below shows a melism with a falling slope on two adverbs “très, très” (about 12 semi-tones) modifying an adjective “chouette”.

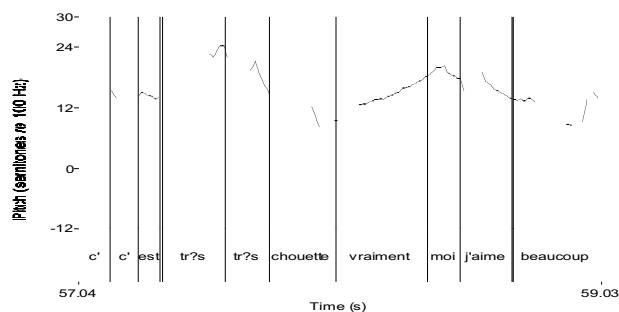


Figure 2: melism for “très très (chouette)”

(iii) to spatial coordinates (names of streets or buildings, directions)

Discursive and phatic exchanges (e.g. “attendez”, “ah remarque”, “figurez-vous”, “hein ?” ...) are quite frequent. Melisms associated with the technical task are quite rare but this is because this task is rather marginal with respect to the other tasks which are more closely linked to the objectives of the tourist.

Figure 3 below shows a sequence of 2 melisms produced on the words “prenez” (with a plateau at 340 Hz) and “souris” (about 8 semitones).

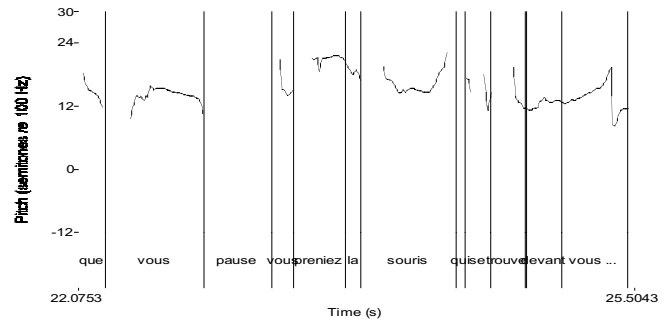


Figure 3: melisms on the words “prenez” and “souris”

For these melisms, we identify three layers of meaning: perlocutory aims, values and functions.

4.2. Subjective melisms and ‘perlocutory aims’

As is well-known, intonation has a number of functions of a morphosyntactic, syntactic or simply grouping nature. Prosody also, however, plays a role of social convention in that it is the area of linguistic structure as social heritage, the milieu of an inherited, learned, objectified and shared practice. It is regulated by the social convention which bestows on it a normative framework (standardisation), exercising an action of a coercitive type, and is therefore the place of recurrent prosodic forms. But by way of consequence, this space allotted to structure, the place of social exercise, leaves little space for subjective expression. Now the message, in order to be received in an optimal manner by a person or a group of persons, must be perceived (and therefore produced) as the expression of a belief. In order to giving rise to this belief, then to adhesion (but also in a non-voluntary manner, its corollary, to refutation) and even to action, the message must carry a subjective dimension, the space of individual encounter. This is necessarily embodied in physical parameters and in a space, by nature and by necessity, distinct from the first.

In this framework, underlying beliefs are analysed as perlocutory aims which take on the dynamics of interaction in the dialogue, mobilising at the same time superficial beliefs more directly linked to the lexicon. These perlocutory aims are related to the objectives of the dialogue and in particular to the instructions given before the recording. The underlying beliefs are the following:

- the need to convince the ‘tourist’ to include cultural visits in his itinerary,
- the desire to help the tourist with a choice of activities, of itineraries, providing information on tourism, the town, gastronomy, technical help with the computer interface, diplomatic dialogue management. This type of believe can be inferred from the role which the speaker assumes or may natural and spontaneous,

These underlying beliefs are actualised on a lexical level by local beliefs applied to different words belonging to the different lexical fields detailed above. The principal local beliefs, whether conscious or not, can be expressed as follows:

- a given tourist object (“château”, “église”, “restaurant traditionnel”, “équateur”, “boutiques”, etc.) is worth a visit to the tourist,
- a given specific quality (“joli”, “dépayçant”, “chics” etc.) of this object is both shareable and motivated,
- a given object of the map or the town (“boulevard Voltaire”, “rond-point”, “centre commercial”, “sens unique”, “première rue” etc.), or a given action or movement (“tournez”, “prenez”, “cliquez” etc.), or a given technical object (“souris” etc.) is a necessary action for the itinerary.

4.3. Melisms et values

In fact these beliefs are based on values which are either inherited from instructions and assumed tasks or which are specific to the speaker. This combination of values is only possible in so far as they are not in contradiction with each other: personal values reinforce, relay, prolong or extend the circumstantial values imposed by the instructions. The principal circumstantial value can be expressed as follows:

- follow the instructions (e.g; establish a programme of cultural visits, update the map, elaborate an itinerary, inform).

The principal values introduced by the speaker are the following:

- insistence on natural elements (the park with its trees, rose-gardens, ponds, birds etc.),
- insistence on traditional elements (market, handcrafts, costumes etc.),
- insistence on human interaction (hospitality and atmosphere of restaurants, friendliness of shopkeepers etc.).

4.4. Melisms and functions

The functions of melisms are naturally very much dependent on the type of dialogue, the objectives and beliefs, the inherited or personal values. In the framework of this corpus limited to the production of a single speaker the values are the following:

- elective function: drawing attention to a given lexical item because of its informative content (touristic, computing, strategic, procedural),
- sharing belief in a value by use of the affective pitch register (High level). This affective register is in particular that of emotions and memories attached to them. Remembering, evocation, suggestion are in this context important manipulatory values. Whether or not they are conscious they can be euphoric or dysphoric, converging or not with those of the listener and leading or not to an action or a reaction.

5. Conclusion

Melisms, in so far as they make use of the High or Extra-High pitch registers are a preferential means of expression of ordinary subjectivity. By use of the high pitch registers and melisms, the speaker selects, whether voluntarily or involuntarily, that information which she considers essential to convey an element of subjective or objective information.

The sharing of the affective register allows communication to be established beyond the range of linguistics between two subjectivities. It is this which instantiates priorities carried by linguistic information. It is the area of inter-individual contact which can lead either to a sharing of values when beliefs are common or compatible or to a confrontation when they are not and which may subsequently result in action. Discourse without melisms is an exterior discourse which avoids contact and subjective communication, a discourse where the individual is in some sense absent.

Acknowledgements

My warm thanks to Daniel Hirst for his help with the English version of this text.

6. References

- [1] Austin J.-L., 1991. *Quand dire, c'est faire*. Tr. par G. Lane. Paris: Seuil, Points.
- [2] Caelen-Haumont, G., 1981. *Structures prosodiques de la phrase énonciative simple et étendue*. Hamburger Phonetische Beiträge, Band 34, Hamburg Buske.
- [3] Caelen-Haumont, G., 2002. Towards Naturalness or the Challenge of Subjectivity. In Keller, E., Bailly, G., Monaghan, A., Terken, J., Huckvale, M. (eds.), *Improvements in Speech Synthesis*. London.: John Wiley.
- [4] Caelen-Haumont, G., 1991. Stratégies des locuteurs en réponse à des consignes de lecture d'un texte: analyse des interactions entre modèles syntaxiques, sémantiques, pragmatique et paramètres prosodiques. Thèse de doctorat d'état, Université de Provence, Aix-en-Provence.
- [5] Caelen-Haumont, G.; Bel, B., 2000. Le caractère spontané dans la parole et le chant improvisés: de la structure intonative au mélisme. *Revue Parole*, 15/16, 250-301.
- [6] Caelen-Haumont, G.; Bel B., 2001. Subjectivité et émotion dans la prosodie de parole et du chant: espace, coordonnées et paramètres. Actes du *Colloque International Emotions, Interactions et Développement*, Grenoble, 141-147.
- [7] Cosnier, J., 1994. *Psychologie des émotions et des sentiments*, Retz, Paris.
- [8] Fairbanks, G.; Hoaglin, L.W., 1941. An Experimental Study of the Duration Characteristics of the Voice during the Expression of Emotion. *Speech Monographs*, 8, 85-90.
- [9] Fairbanks G.; Pronovost, W., 1939. *Speech Monographs*, 6, 87-104.
- [10] Fonagy, I; Sap J., 1977. Traits prosodiques distinctifs de certaines attitudes intellectuelles et émotives. Actes des *8èmes JEP*, Aix-en-Provence, 237-246.
- [11] Léon, P.R., 1976. De l'analyse psychologique à la catégorisation auditive et acoustique des émotions dans la parole. *Journal de Psychologie*, 3-4, 305-324.
- [12] Rossi M.; Di Cristo A; Hirst D.; Martin P.; Nishinuma Y., 1981. *L'intonation de l'acoustique à la sémantique*. Paris: Klincksieck.
- [13] Terken, J.M.B., 1991. Production and Perception of Prosodic Prominence. Actes du *12ème ICPhS*, Aix-en-Provence, Vol. 1, 288-293.