How does fundamental frequency correlate with perceived politeness in greetings ?

Isabelle DUMAS[*],Emmanuel FERRAGNE[**]

Abstract

The purpose of this paper is to consider the contribution of prosody to discourse analysis in interaction by means of a concrete example, i.e., the analysis of *bonjour*, a French greeting. In this experiment, naïve subjects were made to listen to a series of French politeness formulae taken from real interactions and asked to evaluate the degree of politeness they perceived. Two hypotheses were tested : we first tried to determine whether the scores of the greetings recorded during commercial encounters were significantly different from those ascribed to greetings extracted from service encounters ; then, leaving aside this dichotomy, we concentrated on a hypothetical correlation between perceived politeness and the F_0 glissando size (in semitones) on the second syllable of *bonjour*.

1. Introduction

Working on a comparative study of service interactions in the fieldwork of *pragmatics* and *discourse analysis in interaction*, we attempted to discriminate between several kinds of service encounters which Aston defines as encounters "where the discourse confirms the basic expectations it sets up for a 'business transaction' as the initial talk-type" (1988:25). This definition raises the problem of the absence of distinction between commercial interactions (e.g., buying newspapers) and service interactions (e.g., asking for information). Actually, clear-cut distinctions between commercial and service situations are not easy to draw. Although the material vs. immaterial, or financial interest vs. no financial interest dichotomies have proved valuable for our previous work, they do not account for all potential problems. Now, what happens at the newsagent's, for instance, where the person selling a product is not the person who produced it ? Since the situational definitions of service encounters do not seem to allow for all cases, we looked for other elements and the linguistic analysis of politeness turned out to be an interesting criterion.

We began to work on the politeness strategies used by the interactants. The theoretical background for this part of the study can be found in the descriptive model of linguistic politeness devised by Brown and Levinson (1987). Their model is based on Goffman's notion of *face* (1971). Each individual possesses two faces : a negative one, roughly corresponding to the territories of the self, and a positive one, roughly corresponding to narcissism. Each time there is an interaction involving two individuals, four faces are brought together. Moreover, all through the interaction, the interactants are led to accomplish a certain number of acts – verbal ones or non-verbal ones – each of them constituting a potential threat for one or several of these four faces. Brown and Levinson called these potential threats *Face Threatening Acts* (FTAs). Their model has been taken up again and expanded by Kerbrat-Orecchioni (1992) who considered that some of these acts can also enhance the same four faces. She calls these acts *Face Flattering Acts* (FFAs).

In the type of interaction we studied, politeness aims at keeping the course of interaction smooth as well as establishing or preserving good relationships between participants. Since politeness is essential in everyday life, a certain number of set formulae are ready for use. The recurrence of one of them – *bonjour* – particularly held our attention. Of course, this politeness formula constitutes a FFA since the speaker flatters the positive face of the hearer by signalling that he has seen him and that he will serve him. Yet, depending on the context in which it occurs or on the speaker who utters it, this greeting can be intuitively perceived by the hearer with varying degrees of politeness.

This is precisely where prosody comes as a complement to better understand linguistic politeness in communicative situations. Indeed the pragmatic analysis of greetings did not allow us to draw any distinction between a *bonjour* perceived as polite by the hearer and another *bonjour* perceived as being more polite than the other by the same hearer.

2. Data

We were led to work from case studies and, more specifically, from empirical data. Five sub-corpora were collected, each of them composed of about forty recordings of a particular type of service encounter according to a

---

[*] Groupe de Recherche sur les Interactions Communicatives – UMR 5612 – Isabelle.Dumas@univ-lyon2.fr

[**] Emmanuel.Ferragne@univ-lyon2.fr

preliminary classification based on the criteria mentioned in the introduction. Each recording corresponds to an interaction and begins with a greeting. The sub-corpora were recorded with a hidden microphone and/or camera. The shopkeepers and the clerks knew they were being recorded while the customers and the users did not. The first two sub-corpora were collected in a commercial situation : at a stationer's and at a newsagent's [1]. An intermediary corpus was collected at a post-office because both commercial (selling of stamps) and service (banking relationship) situations occur. The last two sub-corpora represent a service situation : first a paying service in a bank and a free one in a town hall.

We took the greetings from the *opening sequences* of the interactions and concentrated on the greetings uttered by the shopkeepers and the clerks. These greetings were submitted to a qualitative and quantitative pragmatic analysis which proved unable to provide any satisfactory explanation for the differences perceived by the hearers. Following Di Cristo's remark (2000:21) concerning the complementarity of prosody and interaction studies, we felt that resorting to prosody had become necessary in order to find out which parameters were responsible for perceived politeness.

3. Procedure

From the beginning, we had to leave out some of our greetings : quite a few of them were found to overlap, and superimposed background noise was fairly frequent. Before preparing the actual stimuli for the experiment we classified the *bonjours* from our corpora according to F_0 criteria. The final classification was obtained after going through four main stages of data reduction. A summary for the various stages of the classification is given here to help understand the present study. We first converted fundamental frequency values – in Hertz – to semitones[2]. This was meant to provide us with relative values and a more manageable unit for further calculation. It is true that other scales are said to be more appropriate for reflecting how people perceive pitch, in particular the Equivalent Rectangular Bandwidth. However, the practical advantages of the semitone scale, plus the fact that our reference work ('t Hart, Collier & Cohen, 1990) used the musical scale, led us to opt for this solution. The next step consisted in spotting octave jumps – which were not infrequent – and correcting them. This can be easily done in *Praat*[3] by setting the argument *Octave-jumps cost* to a different value, or using the *Kill octave jumps* function. The resulting pitch curves were then stylized so as to erase the variations that were not linguistically pertinent. We finally designed a procedure intended to help decide whether a given F_0 glide was perceptually relevant or not. F_0 glides from the first and second syllable of all *bonjours* were extracted from the original signal. A *Praat* script was written to compute automatically glissando duration, glissando size, glissando rate, maximum intensity, and indicate perceptual relevance using the formula $y = (400/x)^2$ where y is the minimum rate required (in semitones per second, ST/s henceforth) for a pitch slope to be perceived for a given x-second glissando[4]. If measured glissando rate was lower than had been calculated with the above formula, the syllable in question was assumed to have a perceptually steady pitch pattern. The results were then pasted to a spreadsheet program. It appeared that the *bonjours* were most frequently (37 per cent of the time) characterized by a pattern consisting of a stationary F_0 on the first syllable and a rising F_0 on the second syllable. The "steady-rising" *bonjour* which exhibited average duration, intensity and F_0 was selected as a basis for creating the stimuli used here.

3.2. Stimuli

From this prototypical *bonjour*, another *bonjour* was generated with a steady pitch pattern on the second syllable. Then the final pitch point on the second syllable (initial absolute value : 330 Hz ; the speaker was a female with a noticeably high-pitched voice) was dragged by successive 1-semitone steps up to 25 semitones away from its original position (the pitch point at the beginning of the syllable remaining unchanged) and down to minus 17. This

---

[1] In order to avoid any misinterpretation due to cultural specificities, what we call "newsagent" is in fact a *tabac-presse*, i.e., a place where you can buy cigarettes and newspapers.

[2] The context is expected to clarify the meaning of the term F_0, referring in this article to both the objective measure of vocal fold vibration as estimated by the pitch algorithm (in Hz) and – perhaps more loosely – the attempt at quantifying perceived pitch (in semitones).

[3] Software for speech analysis developed by Paul Boersma.

[4] The formula is derived from 't Hart and colleagues' findings that "[in] order to be audibly distinguished from a stationary tone, a short glide of 50 ms duration should have a rate of change of F_0 (g) of sixty-four semitones per second (ST/s). Each doubling of the duration decreases the threshold value by a factor of four."('t Hart, Collier & Cohen, 1990:32)

operation yielded a set of 43 manipulated *bonjours* which, after an auditory assessment, was reduced to 26 because the members lying outside the [-8, +17] interval sounded unnatural. The stimuli for the other part of the experiment were 4 *bonjours*, 2 *mercis* (thank you), and 2 *au revoirs* (good bye) taken from each site. The 66 stimuli (26 + 40) were randomized, separated by a 5-second silent interval and stored as one single sound file. On account of between-speaker differences in sound power and further intensity variations due to uncontrolled distance between speaker and microphone (during the recording of the corpus), and the acoustic characteristics of each site, we had to re-adjust intensity levels so that the whole set would exhibit approximately homogeneous loudness. It was expected that having the subjects assess the stimuli from the two experiments as if there were only one would play down any possible learning effect and make the task less tedious.

3.3. Experiment

The 37 subjects were university staff and students. They were asked to give a mark out of twenty for the politeness they perceived in each word ; 0 meaning *peu poli* (not very polite) and 20 *très poli* (very polite). We could have used a Likert scale with, say, five points and as many explicit labels associated with them ; the reason why we went for the traditional mark out of twenty is because we wanted to examine how people would rate politeness with reference to a scale they had been familiar with since their childhood. The subjects were told to do their best to disregard background noise for their evaluation. The experiment actually started with a short practice test the aim of which was to allow the subjects to get a better idea of the conditions and to make sure everybody could hear the stimuli.

4. Results and discussion

We used the Mann-Whitney rank test to examine the following hypothesis ; H_0 : on average, politeness in commercial interactions is given the same score as politeness in non-commercial interactions. The values from the post-office were not included for the calculation of the statistic because it was not possible to decide whether it belonged to the commercial or non-commercial type. The value for the test-statistic – U = 71 – turned out to be significant at the 5 per cent level (the critical value for $p < .05$ with $n_1 = n_2 = 16$ is 75). In other words, the subjects seem to have perceived more politeness in formulae taken from commercial situations. Note however that the interval between the value required to achieve significance and U is far from large ; it would therefore be safer to redo the experiment with bigger samples from our corpus.

However, these results suggest that new elements are available for the classification of interactions in service encounters. It seems, indeed, that politeness is perceived differently whether its source be, on the one hand, the stationer's and the newsagent's, and, on the other hand, the bank and the to wn hall. On the face of it, we might believe that shopkeepers use a higher degree of politeness since the outcome of the interaction involves financial profit (Dumas & Ferragne, 2003). Shopkeepers are thought to make extensive use of what we may term "selling greetings" which come complete with a smile[5]. This is an oversimplification : we could argue that town hall clerks, for instance, adopt an authentic type of politeness simply because they are constantly polite ("[*je suis*] *à votre service*")[6] despite there being no financial interest at stake. Another factor can be introduced to account for the difference revealed by statistical analysis, namely, the nature of the relationship between participants. As far as the newsagent's and the stationer's are concerned, their products are basic consumables, therefore, most customers are regular ones. A relationship is likely to develop between the shopkeepers and the customers and this may be shown by the addition of a prosodic parameter[7] as a means employed by the shopkeepers to manifest their accepting this social relation. The situation at the town hall and the bank is the reverse and it seems to lack this prosodic parameter that belongs to interactions where people know one another rather well. Granted that *bonjour* is – semantically – an FFA and given that the added prosodic parameter at the stationer's and the newsagent's may be considered as another FFA (since the subjects did perceive a higher degree of politeness there), it may be claimed that shopkeepers use double FFAs.

---

[5] Note, incidentally, that a smile has both visual and acoustic cues ; the acoustic correlates seem to be detectable in the location of the first three formants (Aubergé & Cathiard, 2003:94-95).

[6] We cannot think of any accurate translation for "(je suis) à votre service". Clerks sometimes use it at the end of interactions to minimize the importance of the service they have just supplied. A rough equivalent would be "don't mention it".

[7] This analysis pre-supposes a binary categorisation – a matter of either/or – i.e., reasoning in terms of the absence or presence of one feature. Although this is quite convenient and acceptable for linguistic analysis, recall that the prosodic parameter mentioned may possess more than two degrees, perceptually speaking.

For the second part of the experiment, the *bonjours* were ranked according to F_0 glide criteria in such a way that the –8 glide ranked 1, the –7 ranked 2, and so on. This objectively defined ranking was compared with the ranks given by the subjects' scores (the lowest score ranked 1st and the highest one ranked 26th). We then computed a Spearman correlation coefficient : *r_s = 0.889*, which is highly significant (at least at the one per cent level : at the time of the experiment, we only had paper versions of statistical table and consequently had to make do with default confidence intervals, which, in itself, remains inconsequential). As can be seen in Figure 1, there is a high positive correlation between F_0 glide and perceived politeness. Rank correlations cannot be interpreted in quite the same way as the Pearson coefficient (see Woods, Fletcher & Hugues, 1986:169-174). In particular, r_s^2 – contrary to (Pearson's) r^2 – does not seem to indicate reliably (at least under certain circumstances) to what extent the observed variation in variable *x* can account for the variation in variable *y*. Suppose r_s^2 were reliable in our case, this would mean that an increase in glissando rate (not in terms of absolute rate, here, the sign matters: e.g., a -4 ST glide is to be considered as lower than a +2 ST one) accounts for about 79 per cent of the corresponding increase in perceived politeness. Whatever the caveats in interpreting this percentage, glancing at Figure 1 soon convinces us that 79 per cent may not be too bold an estimate. A more pessimistic explanation would be that some subjects managed to guess the logic behind the experiment and focused on pitch glides for their judgments. We believe that the layman is quite capable of seeing through what we expected and inferring, on common sense grounds, that we were interested in studying the relationship between pitch and perceived politeness. However, although this bias may have led us to overestimate the correlation a little, it is very unlikely that a great number of subjects actually made this guess ; therefore, we may safely conclude that we have found out a true high positive correlation.
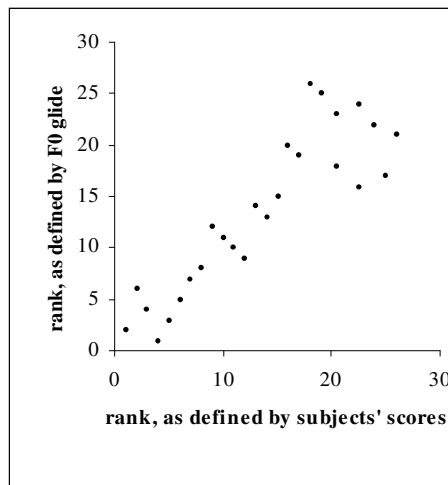

Figure 1 : A scattergram displaying the correlation between perceived politeness and glissando size
(the -8 ST glide gets rank 1, the -7 rank 2, etc., while the lowest score gets rank 1)

It should also be noted that most subjects laughed when hearing the manipulated stimuli with the +15 to +17 glissandos. Such sizable glissandos are hardly ever heard in everyday speech, not to mention the fact that the absolute values would imply an unusually high register ; the last third of the pitch curve in such cases was above the 600 Hz ceiling generally accepted as the upper limit for F_0 in speech. We may also wonder what property of a given glissando is relevant for the perception of politeness and naturalness : size, or absolute starting and finishing values, or slope. It is most probable that the subjects interpreted the final parts of such glissandos as non-human squeaks, or flaws that had appeared during the editing process. It remains to be seen why the *bonjour* that scored highest (i.e., the +12 ST stimulus) also had a glide rising above the 600 Hz ceiling. A tentative answer is suggested by Rossi (1978:14) : "un glissando n'est jamais perçu dans sa totalité, sa hauteur est déterminée par la fréquence située entre le 2e et le 3e tiers de la durée vocalique". Therefore, it may well be that the final third of the glissando was not perceived. We may now close this section on the acoustic properties of our optimally polite *bonjour* by providing values for the pitch curve on its second syllable. The glissando for this *bonjour* is, as we have just said, a 12 ST rising one, i.e., the final frequency is obtained by doubling the original frequency (in Hz). The glissando is 185 ms long, hence a slope of approximately 64

ST/s. This last value must be re-utilized with caution because it may be that a change in absolute duration would require that the slope be re-adjusted (recall that the perception of a pitch glide is affected by its duration, see footnote 4). We may conjecture that it would be possible to define the optimally polite F_0 slope (all other things being equal) with reference to the minimum slope required for the contour to be perceived as a glide for a given duration (see formula in 3.1).

Rank	Stimulus	Rank	Stimulus
1	12	14	5
2	8	15	0
3	13	16	1
4,5	7	17	2
4,5	15	18	3
6,5	9	19	-1
6,5	14	20	-2
8	16	21	-4
9	17	22	-6
10	10	23	-8
11	11	24	-5
12	6	25	-3
13	4	26	-7

Table 1 : *Bonjours* **sorted by rank**

(Each stimulus is identified by the size of its final glissandos ; the most polite one ranks first, as opposed to Figure 1)

From the point of view of pragmatics, this correlation seems to emerge from the fact that the final rising glide on the second syllable of *bonjour* actually comprises two illocutionary forces (Austin, 1962). Its primary force is contained in the semantic content of the word itself ; the secondary force is conveyed by suprasegmental features : it is an interrogative value. The course of all interactions is anchored in a script, that is "a predetermined stereotyped sequence of actions that defines a well-know situation" (Schank & Abelson, 1977:41). The script in service encounters stipulates that after exchanging greetings with the shopkeeper or the clerk, the customer or the user must express his request, which is a delicate matter since it constitutes a FTA. The request can threaten the hearer's negative face if s/he thinks it might deprive her/him of part of her/his freedom of action, even though the threat is somewhat weakened in service encounters because it is necessary for the script to follow its normal course (Dumas, 1999:133). With the final rising F_0, the shopkeeper (or clerk) adds a secondary value of asking for the customer's request which can be paraphrased as "what would you like ?". This glide may be interpreted as a FFA from the shopkeeper (or clerk) intended to attenuate the FTA the customer is about to produce and to incite him to make his request without further ado. It is worth noticing that, once again, the stationer's and the newsagent's can be singled out from the rest owing to their great percentage of "steady-rising" *bonjour* : 79 per cent at the stationer's, 68 per cent at the newsagent's, 47 per cent at the town hall, 25 per cent at the post office and 22 per cent at the bank. The conclusions here match what has been discussed previously about the special status of the stationer's and the newsagent's.

5. Conclusion

Having reviewed the results of our study, we may now discuss the usefulness of combining a prosodic and pragmatic approach as employed in this study. At the outset, the approach was a top-down one, i.e., starting with a pragmatic model to try and assess to what extent the model would fit empirical evidence. To this aim, we first used pragmatics as a tool for the inspection of the data, this time, in a bottom-up process, that is working from raw data to abstractions. Here, just as in any study involving both theoretical and practical aspects of a given discipline, we intended to estimate the degree of matching found when the "bottom" part (to keep our conclusion very visual) eventually meets the "top" one. As it turned out, the fact that empirical data require accurate gauging of continuous (as opposed to categorical) entities made pragmatics somewhat ineffective for the scrutiny of this kind of data. It is prosody that proved most successful in dealing with the bottom-up part and providing us with adequate tools for observation as well as experimentation. To be more specific, it is what we may call *prosodetics* (as opposed to *prosodemics*) that was used as the empirical counterpart of pragmatics. It is hoped that an increasing amount of research will keep moving towards an integrated view of language through the cooperation of researchers from several fields of investigation.

Bibliographical references

Aubergé V. & Cathiard M., 2003, "Can we hear the prosody of smile ?", Speech Communication 40, p. 87-97.

Austin J. L., 1962, *How to do things with words*, Oxford, Oxford University Press.

Aston G. (ed.), 1988, *Negotiating service*, Bologne, CLUEB.

Brown P. & Levinson S., 1987, *Politeness. Some universals in language usage*, Cambridge, Cambridge University Press.

Di Cristo A., (forthcoming), "Interpréter la prosodie", in Actes des XXIIIèmes Journées d'Étude sur la Parole, Aussois, 19-23 juin 2000.

Dumas I., 1999, *Les interactions verbales en situation de commerce : étude comparative réalisée dans une librairie-papeterie-presse et un tabac-presse*, Mémoire de DEA de sciences du langage, Université Lumière, Lyon 2.

Dumas I. & Ferragne E., 2003 (in press), "La prosodie : un marqueur de politesse", in Actes des Journées Prosodie 2001, Grenoble, 10-11 octobre 2001.

Goffman E., 1971, *Relations in public*, New-York, Basic books.

't Hart J., Collier, R. & Cohen, A., 1990, *A perceptual study of intonation : an experimental-phonetic approach to speech melody*, Cambridge, Cambridge University Press.

Kerbrat-Orecchioni C., 1992, *Les interactions verbales, tome II*, Paris, Armand Colin.

Rossi M., 1978, "La perception des glissandos descendants dans les contours prosodiques", Phonetica 35, p. 11-40.

Schank R. C. & Abelson R. P., 1977, *Scripts, plans, goals and understanding : an inquiry into human knowledge structure*, Hillsdate, Lawrence Erlbaum Associates.

Woods A., Fletcher P. & Hughes A., 1986, *Statistics in language studies*, Cambridge, Cambridge University Press.