CONVERGENCE OF FUNDAMENTAL FREQUENCIES IN CONVERSATION: IF IT HAPPENS, DOES IT MATTER?

Belinda Collins
Phonetics Laboratory, Department of Linguistics (Arts), Australian National University

ABSTRACT

This paper explores the existence and nature of accommodation processes within conversation, particularly convergence of fundamental frequency (Fo) of conversational participants over time.

The study raises a number of issues related to methodologies for analysing interactional (typically conversational) data. Most important is the issue of the applicability of statistical sampling methods which are independent of the interactional events occurring within the talk. It concludes with suggestions for a methodology that examines long term acoustic phenomena (eg long term fundamental frequency) and relates events at the micro acoustic level to interactional events within a conversation.

1. INTRODUCTION

This paper discusses a pilot study which looks at the relationship between phonetic realisation and social processes. It arose out of an interest in Conversation Analysis and sociolinguistic readings in Accommodation Theory.

Developments in Linguistics and Sociology in the 1990's, particularly in Conversation Analysis (CA) suggest a view of language use that is highly interactive. The view is that language and its component parts "*is designed for interactional ends and as such must reckon with the arthitecture and dymanics of turns, sequences, activities, ... and other relevant interactional actualities*" (Schegloff, Ochs & Thompson 1996:36). If this is so then it is possible that the effects of interaction are manifest at all levels of linguistic production, including at the level of phonetic production.

This view has been supported by sociolinguistic research being undertaken from a social psychological perspective. Giles' Accommodation Theory focuses on the interactive aspects of interpersonal communication and has identified the occurrence of 'convergence', that is "*the processes whereby individuals shift their speech styles to become more like that of those with whom they are interacting*" (Giles & Smith 1979:46). Convergence in pronunciation, speech rates, pause and utterance lengths, and vocal intensities have all been observed, as has convergence of the fundamental frequency (Fo) of the babbling of infants in the presence of their mother or father (Giles & Smith 1979:46).

Consequently there are indications that intonation is not only a phonological resource (Ladd 1996), and a means whereby text structure is indicated (Halliday 1967, Brown 1983), and turn exchange opportunities signalled (Sacks, Schegloff & Jefferson 1974; Selting 1998), but also a resource for demonstrating larger social processes.

Gregory (1991) states that "*Spectral analysis of the fundamental frequency band in speech of interview partners reveals covariance of voice energy levels, and thus, a possible form of rudimentary social synchrony*". The aim of the experiment under discussion here was to determine whether the mean Fo of conversational partners would converge over time and whether there were any other relationships between the Fo of conversational partners over time. However during the course of the analysis it became obvious that there were problems associating movements in Fo levels with 'rudimentary social synchrony'. The later part of this paper will discuss the pitfalls inherent in linking decontextualised observations to context embedded social events.

2. METHODOLOGY

The subjects of the experiment were eight women, paired according to age, education and other demographic factors. Conversations took place within a recording studio in order to optimise sound quality. No topic was set for the conversations and there was no particular task to complete. This was to avoid the consequences that prestructuring of the interaction may have on the nature of the conversation and to minimise the effect of the setting and the experiment on the content of the talk. (Schegloff, Ochs & Thompson 1996:21). Participants were paired carefully to avoid unequal social relations as there is some evidence that this influences who accommodates and to what degree (Gregory 1991), and to make sure that they had enough in common so that they would be able to talk to each other for fifteen minutes without running out of things to say.

Participants were first asked to read two lines of a poem. This enabled the taking of a decontextualised baseline measurement. They were then left to talk for approximately fifteen minutes.

The recordings were digitised using the program Signalyse. An FFT-Comb analysis was undertaken for each sample, with a frequency range setting of 100-400Hz, a new spectrum every 5 ms, a threshold of 20% and an output filter setting at 8 extractions. The sample duration was constrained by both

the technology and the nature of natural conversation. The Signalyse programme was most comfortable sampling portions of three to four seconds in duration. Longer samples resulted in extremely slow processing and in many cases the application quitting altogether. Nevertheless, in three to four seconds of conversation, speakers can exchange turns six times. They can have significant portions of overlap, silence. laughing, loud breathing sniffing, coughing and so on.

Another characteristic of natural conversation in English is that it is typically linear. One speaker speaks at a time. Sampling of two speakers cannot be done at once. After the baseline sample was taken the conversation was sampled at roughly one minute intervals nine times. A piece of talk that was free of turn exchange and extraneous noise was sampled for three seconds and a portion typically between 1 and 1.5 seconds in duration selected and subjected to analysis to extract the mean fundamental frequency. The same procedure was then carried out on a subsequent portion of talk produced by the alternate speaker which was as close as possible to the first. This procedure resulted in ten pairs of mean Fo figures. each sampled over 1 to 1.5 seconds.

The data was tabulated, line plots produced and correlation tests carried out to test a number of hypotheses.

3. RESULTS

The hypothesis that the mean Fo of speakers would converge at some point in a two party conversation is supported by the data.

The observations showed that the mean Fo of participants in a two party conversation did converge at a number of points throughout the portion measured in that the differences between the mean Fo for each speaker at these points is smaller than the difference in their baseline measurements. However, once achieved this convergence was not maintained by participants.

In addition. some speakers appeared to be maintaining their mean Fo so that a relationship in the Fo readings of the two participants was maintained over a number of sample points. Gregory's finding of covariance was supported.

As shown in Table 1 and Figure 1, for Pair One the difference in Mean Fo at the baseline was 29.25Hz. The Mean Fo converges. ie: becomes less that 29.25 Hz at sample points 3 and 9. The mean Fo of participants does covary across three points. The relationship between mean Fo is maintained (within 5Hz) over consecutive sample points at points 4, 5 and 6.

Sample	Sp 1 Mean Fo	Sp 2 Mean Fo	Difference Mean Fo Speaker 1 - Speaker 2
baseline	190.57	161.32	29.25
2	236.53	182.43	54.10
3	185.65	176.07	9.58
4	197.85	156.04	41.81
5	220.11	177.71	42.40
6	194.92	157.62	37.25
7	241.46	160.64	80.82
8	230.98	167.20	63.78
9	158.24	170.90	12.66
10	267.92	183.50	84.42

Table 1: Pair One, Mean Fo results plus difference between speakers in Mean Fo levels at each sample point.

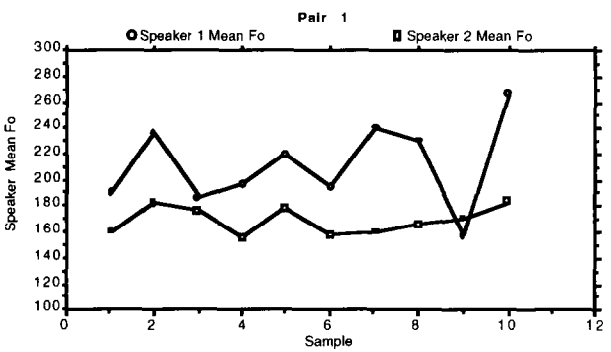

Figure 1: Pair One, mean Fo Samples 1-10

As shown in Table 2 and Figure 2, for Pair Two a significant difference in mean Fo was not apparent. At no subsequent point in the data set is the difference in Mean Fo equal to or less than the baseline reading of 2.81Hz. The mean Fo of participants does not converge but does covary at two points. The Mean Fo at data points 3 and 4 is equidistant (within 5 Hz).

Sample	Speaker 1 Mean Fo	Speaker 2 Mean Fo	Difference mean Fo Speaker 1 - 2
Baseline	196.38	193.57	2.81
2	167.79	179.26	8.47
3	193.09	227.09	34.00
4	176.78	207.40	30.62
5	229.44	220.62	8.82
6	190.09	224.42	34.33
7	205.16	199.72	5.44
8	188.22	193.58	5.36
9	179.45	246.21	66.76
10	184.93	210.00	25.07

Table 2: Pair Two Mean Fo results plus difference between speakers at each sample point.

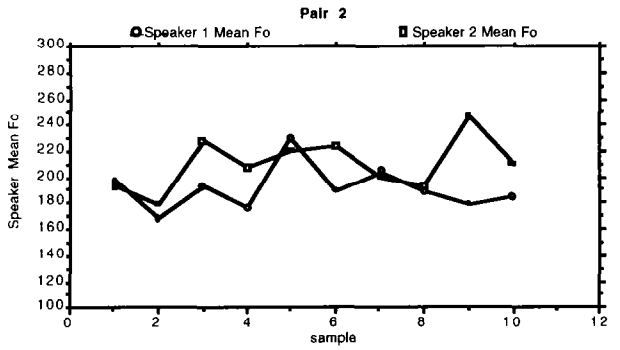

Figure 2: Pair Two Mean Fo Samples 1-10

As shown in Table 3 and Figure 3, for Pair Three the difference in Mean Fo at the baseline is 10.95 Hz. The Mean Fo of the two speakers converge at sample point 6 and 7. The mean Fo also covaries at these two points as it is equidistant (within 5Hz).

Sample	Speaker 1 Mean Fo	Speaker 2 Mean Fo	Difference mean Fo Speaker 1-2
Baseline	173.85	184.8	10.95
2	177.57	160.48	17.09
3	178.65	213.36	34.68
4	154.16	178.07	23.91
5	164.24	182.60	18.36
6	158.45	161.05	2.60
7	152.65	159.51	6.86
8	162.65	150.24	12.41
9	153.89	176.81	22.92
10	158.56	166.17	7.61

Table 3: Pair Three Mean Fo results plus difference between speakers at each sample point.

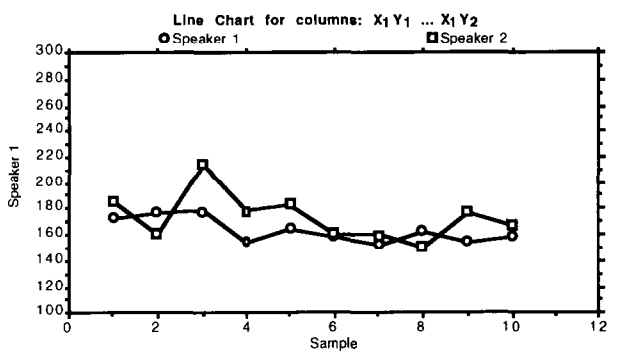

Figure 3: Pair Three Mean Fo Samples 1-10

As shown in Table 4 and Figure 4, for Pair 4 the difference in Mean Fo at the baseline is 41.31Hz. The Mean Fo of the two speakers converge at sample points 2, 4, 6, 7, 8 and 9. At these points the difference in Mean Fo is less than the difference at the baseline. However the convergence is not maintained at a consistent or level across these points, nor is the convergence intensifying (the difference decreasing at each point). Rather whilst the difference at these points may be less that the difference at the baseline, the relationship between Fo levels at these points does fluctuate.

Sample	Speaker 1 Mean Fo	Speaker 2 Mean Fo	Difference mean Fo speaker 1 -2
baseline	196.17	30.89	41.31
2	152.47	36.10	6.91
3	262.30	74.81	96.11
4	146.68	27.75	3.6
5	167.39	36.57	55.37
6	204.46	72.30	8.09
7	170.45	51.36	12.59
8	220.75	70.81	20.45
9	152.64	13.79	11.57
10	229.95	60.81	67.07

Table 4: Mean Fo results plus difference between speakers at each sample point.

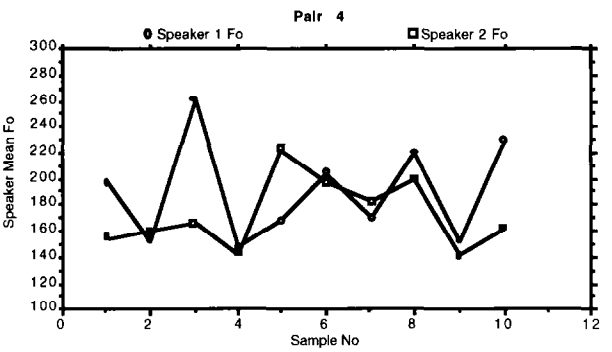

Figure 4: Pair Four Mean Fo Samples 1-10

4. DISCUSSION

The experiment outlined here does confirm that the mean fundamental frequency of conversational participants will converge at some points within an interaction, and will covary. However, the question of whether participants in a conversation modify their Fo in order to establish and maintain a relationship between their mean Fo that that of their conversational partners is a much more complex question. It is not possible to answer this question from the data for a number of reasons.

Firstly, the samples were decontextualised. No consideration was taken of the linguistic content of the utterance, nor the conversational (social) action the talk was produced to achieve. For example, at the linguistic level, without reference to context and content we cannot determine whether intonation patterns that are said to be indicative of particular speech acts such as 'question', 'demand', 'pleas' etc. (Couper Kuhlen 1985158-172) were consequential to the observations. Nor can we determine whether the portion of talk under examination was being

produced to achieve affiliation or disaffiliation at a social/interactional level. Claims that social synchrony are being demonstrated at the level of acoustic phenomena would be stronger if they were supported by evidence of synchrony at other levels of analysis.

Secondly, claims as to the macro social nature of any piece of data are difficult to substantiate if it cannot be shown that the features identified by the analyst are also elements of the talk being observably oriented to by the participants. In other words, can we say that the convergence identified in this experiment is apparent to and relevant to the participants and being used by them as an interactional resource (Selting 1998). If we cannot show this, then we cannot claim more than statistical coincidence.

For these reasons, whilst it is possible to see that the mean Fo of participants in a conversation does converge and covary, claims that this demonstrates accommodation or a form of social synchrony are not sustainable.

5. Conclusion

The solution to these methodological and analytic problems is to combine instrumental analysis with detailed transcriptions. This would allow concurrent examination of acoustic, linguistic and interactional content. It would make possible the correlation of acoustic events (made evident by the instrumental analysis) with the linguistic content and the social actions of conversational participants. With this approach, it should be possible to show that relationships in events at the acoustic level are related to and supported by linguistic and social events, and that they are relevant to the participants in the interaction and not just interesting analytical constructs.

Further examination of the conversations produced by the participants in this study are currently being undertaken by the author using the suggested methodology.

6. REFERENCES

1. Brown, G. "Prosodic Structure and the Given/New Distinction" in A. Cutler & D.R. Ladd (eds) *Prosody: Models and Measurements.* Springer-Verlag Berlin Heidelberg (1983)

2. Couper-Kuhlen, E. *An Introduction to English Prosody.* Niemeyer. Tubingen (1985)

3. Giles, H. & Smith, P., "Accommodation Theory: Optimal Levels of Convergence" in Giles, H. & St Clair, R., (eds) *Language and Social Psychology.* Basil Blackwell, Oxford, (1979).

4. Gregory, S.W., *Fast Fourier Transform Low Frequency Spectral Analysis of Mutual Adaptation in Interview Partner's Speech* American Sociological Association (1991)

5. Gregory, S.W., "Sounds of Power and Deference: acoustic analysis of macro social constraints on micro interaction" in *Sociological Perspectives*: 1994, 27.4 winter 497-526

6. Halliday, M.A.K., *A Course in Spoken English: Intonation* Oxford University Press, Oxford (1967)

7. Ladd, D.R., *Intonational Phonology.* Cambridge University Press, Cambridge (1996)

8. Sacks, H., Schegloff, E.M. & Jefferson, G. "A Simplest Systematics for the Organisation of Turn Taking in Conversation" in *Language* 50:696-735. (1974)

9. Selting, M., *TCUs and TRPs: The Construction of Units in Conversational Talk.* InList (Interaction and Linguistic Structures, prepublication web site) http://www.lin.uni-konstanz.de/home/couperku/inlist-index.html

10. Schegloff, E.M., Ochs, E., & Thompson, S.A., "Introduction" in *Interaction and Grammar.* Cambridge University Press, Cambridge (1996)