BRIDGING THE PHONOLOGY - PHONETICS GAP — A METATHEORY

N. Sifakis

University of Essex, Department of Language and Linguistics, Colchester

0. INTRODUCTION

The world of linguistics has seen, over the years, a substantial number of changes, all of them obeying current philosophical, psychological and engineering principles. It has been common practice to distinguish between pre- and post-Chomskyan linguistics, not least because of the radical philosophical and methodological changes that were brought about by the so-called 'Chomskyan revolution'. However, it is noted that Chomsky and his followers concentrated on a purely mentalistic account of language, working essentially on figuring out certain 'universal' principles underlying language competence and usage, under the general dogma of nativism. Apart from a number of production rules permeating production and perception (in the paradigm of 'analysis by synthesis', see Stevens & Halle 1967), there is little, if any, mention of specific analytic ways in which human beings 'translate' thoughts into acoustic signals and vice versa. It is characteristic that Chomskyan linguistics has virtually excluded phonetics from its research, presenting phonology as the output of thought encoding (Chomsky 1972).

If the exclusion of phonetics is characteristic of transformational generative linguistics, the need to create a phonology which could be appropriately articulated has been very strong among more recent so-called non-linear attempts. As we will see below, however, all attempts to conjoin phonology and phonetics can be seen as side-effects of research aiming at resolving problems more central to specific phonological or phonetic theories, rather than genuine effects of a more general, unified theory of phonetics and phonology. This has, as a result, brought about the theoretical estrangement of phonetic science from the 'rest of' linguistics, one which, I would like to argue, is due to deeper differences between the phonetic science, on one hand, and the 'linguistic science', on the other, differences of a metatheoretical nature. After we have seen what these differences are (the term 'phonology - phonetics gap' stands for the problem itself), we will introduce notions which attempt to attack this metatheoretical problem directly, and offer reasons why this should be done this way.

[But first, a preliminary note. The term 'metatheory' accounts for the basic, more often than not implicit, predispositionary 'beliefs' that permeate a work of science and/or philosophy. Although these beliefs are never made explicit (some of their less antagonistic followers are not even aware of them), they nevertheless lie there, behind the lines, initially propounded by the revolutionaries who discarded older metatheories, while, as time was going on, being forgotten or misunderstood by 'current practitioners' (cf. Kuhn 1970). Two examples of such metatheories are empiricism and behaviourism (Pavlov 1927) and cognitive psychology or metaphysicalism (Neisser 1967).

PHONOLOGY AND PHONETICS – A METATHEORY

More relevant to this paper, metatheories set the bases for the structuring of theories. For example, if we view the phonology - phonetics 'gap' as another instance of the notorious mind - body problem, we will be confronted with two, possibly three, architectures in which to work. We can take a mentalist (cf. Chomsky *ibid.*), a neuro-biological (cf. Caplan 1987: 432*f.* and references there) or even a radical behaviourist approach, in which case we disregard any 'gaps' by adopting a **monist** metatheory (however, we would not achieve much; this is why Chomsky has attempted to bring in biology to account for his universal grammar innateness hypothesis -see Chomsky 1980). Alternatively, we could combine mind and body by adopting a **dualist** hyper-structure, in which case we would have to tackle the 'gap' directly since it is explicitly there; or we could even be clever and adopt a **functional** architecture, in which case we would view things causally, talking about 'how the stuff is put together' rather than about what it is made of (Fodor 1981: 124; again we would be talking about the results, not the causations, i.e. the 'gap', themselves). This paper is not going to argue for a new metatheory (although it will be made clear that an exclusive 'gap'-oriented one is needed); we will just review a number of existing ones concerning the 'phonology - phonetics gap' problem and discuss some arrangements that could be made.]

This, then, is our outline: first, we take a closer look at the problem itself; secondly, a review of a number of phonetics- and phonology-oriented accounts of the problem is made and an evaluation is attempted; thirdly, a cognitive phonetic metatheory is suggested; and the paper concludes with a number of metatheoretical constraints that every theory attempting to bridge the phonology - phonetics gap should satisfy.

1. THE 'GAP' PROBLEM

Phonology and phonetics differ in many respects: the former deals with the psychological 'reality' and processing of the linguistically accepted set of sounds in languages, the latter tackles the acoustic, neurophysiological, and biological nature and origins of 'uttered' sounds; the former is constrained with reference to prosodic (i.e. suprasegmental), sociolinguistic even, and phonetic criteria and potentialities, the latter is constrained with respect to mechano-inertial and temporal mechanisms; while phonology entertains processing of phonemic strings in notional time, phonetics involves processing of so-called 'phonetic gestures' in real time (cf. the early critical work of the action theorists, notably Fowler (1977 and 1980). One can think of other differences.

Nevertheless, both phonology and phonetics **do** seem to be conjoined somehow: phonologists (Chomsky & Halle 1968: viii and 9, Hawkins 1984: 31-2 and his 'realization rules') and phoneticians (Tatham, 1990) agree that phonology 'assigns' values and phonetics 'implements' the parameters of these values (however, both disagree with respect to the extent that these values are assigned automatically or not). This is where the 'gap' lies. There certainly is some relation-

ship between phonology and phonetics, except that it as yet has not been made clear by the metatheories of either phonology or phonetics just how this gap is to be filled.

It should be made clear that this 'gap' is strictly and only metatheoretical. That human beings are capable of communicating via language is an *ipso facto* proof of the account that current linguistic 'gap'-oriented metatheories are either inadequate or begging the question. Similarly, it would be metatheoretically unsound to equate the correlations and regularities that physics studies and measures with the explanations that only conceptual models (metatheories), which describe events in generalised forms, are able to give. Doing this, of course, is a subject in itself. For now, suffice it to say that both cognitive scientists (Fodor 1983: 101-104) and neurolinguists (Caplan 1987) converge in believing that the gap exists; I would add that it is not 'real' but 'metatheoretical'. It therefore follows that metatheories (or theory architectures) should provide the basis for a direct mapping between theorising about (i.e. describing) the real world and the real world itself.

2. SOME PHONETIC AND PHONOLOGICAL ACCOUNTS

(i) *Phonetics.* What permeates phonetic theorising is the concentration in resolving the phonological problem of feature extraction from a concrete, experimental phonetic point of view. For example, Ladefoged's linguistic phonetics is seen as a phonetic attempt to better concurrent feature theories (Jacobson, Fant & Halle 1952 and Chomsky & Halle 1968), criticised for specifying one-to-one relationships relating distinctive feature matrices to particular phonetic sounds, from the methodological point of view of experimental phonetics. No attempt to view phonetics outside the generative linguistic paradigm and its metatheoretical consequences is made (Ladefoged 1967: 3), phonetics automatically implementing phonological instructions. This is also the case for the motor theory of speech perception (Liberman, Cooper Shankweiler & Studdert-Kennedy 1967, Galunov & Chistovich 1966), according to which the acoustic signal is interpreted by the listener's knowledge of the speaker's articulatory intentions as evidenced by physiological instructions to the vocal apparatus. More recent versions seeing perception as being modular and neurally based (Liberman & Mattingly 1985), the motor theory is not equipped for bridging the phonology - phonetics 'gap', since it does not address the crucial question of the phonology output to phonetics or the phonological acquisition issue.

If the reduction of the metatheoretical 'gap' problem between phonology and phonetics to the mere methodological scruple of inventing the correct linguistic/'phonological' definitions on the basis of physically testable 'phonetic' constraints is the major concern of early phonetic attempts, current work is centred towards explanatory remarks about the objects of speech perception, rather than conjoining an acceptable phonology to current phonetics in a metatheoretically sound manner. This is certainly true of the action theory of speech production (Fowler, *op.cit.*, Fowler, Rubin, Remez & Turvey 1980), which is primarily concerned with the investigation of 'low tiers',

PHONOLOGY AND PHONETICS – A METATHEORY

i.e. the articulatorily realised 'phonetic gestures', their physiological architecture ('coordinative structures') and operations (for example, their degrees of freedom of movement, etc.). Discarding anything cognitive that can be explained away physically, but implicitly working within a generally accepted linguistic infrastructure (Fowler 1990), action theory offers a wealth of physiological measurements, but losing sight of a unified account of phonology and phonetics. Similarly, other phonetic theories studying the acoustically salient properties of the signal (**quantal theory** Stevens 1989) or the ecologically justified sufficiency of acoustic contrast (**theory of adaptive dispersion**, Lindblom, MacNeilage & Studdert-Kennedy forthcoming), however good in proposing solutions to many speech problems, are not adequate enough for attacking the phonology - phonetics 'gap' itself; no unified notation is proposed and no discussion of phonological output and phonetic input is made. From my perspective, this is no surprise: their metatheories are far too restricted for such a thing.

(ii) *Phonology.* It has to be accepted that modern phonology is not what it used to be. Current theoretical (non-linear) attempts concentrate in suprasegmental (syllable, foot, tone group) structures and their relation to intrasegmental ones (**autosegmental phonology**, see Goldsmith 1979) and limit their representations to string-like arrangements of segments and boundaries in order to appropriately deal with stress and pitch phenomena (**metrical phonology**, see Giegerich 1985 and **moraic phonology**, see Hayes 1989). Yet, although all share the belief that phonological representations need to be much more articulated than traditionally assumed and offer an excellent notation to work with (**dependency phonology**, see Anderson & Ewen 1987), the old precondition of *translation theories* that 'there is no problem in deriving a phonetic representation from a phonemic one, and vice versa' (Lass 1984: 19) is implicitly reiterated. Even the recently advanced **multi-tiered phonology** (McCarthy 1988), equipped with phrase boundaries, stress emphasis etc. specifications, could only be *opaquely* mapped, according to one of its proponentns (Fujimura 1991: 28), onto some (supposedly, phonetics-oriented) motor-theoretic timing values, and the same is true for the action-theory-based **articulatory phonology** (Browman & Goldstein 1987), which has scant notational characteristics. In result, one cannot help but feel that current 'linguistics'-oriented metatheories are not very helpfull with respect to bridging the phonology - phonetics 'gap': they are too restrictive. What is needed is a more general framework, in which intra-linguistic (e.g. phonological accounts of suprasegmental strings) and extra-linguistic (e.g. active processes whereby toddlers acquire phonemic oppositions through social interaction [Schachter 1979: 5]) aspects of speech can be readily accounted for.

3. INTEGRATING A COGNITIVE PHONETIC METATHEORY

In consequence, the first thing we should be careful about, working in a cognitive phonetic architecture, is, as it were, the 'backdrop' of the theory. It would be unwise to use the term 'linguistic' with cognitive phonetics, for reasons that I have mentioned. Rather, it would be more appropriate if we substituted 'linguistic' with 'cognitive', and defined the latter in terms of

PHONOLOGY AND PHONETICS – A METATHEORY

whatever involves **voluntary operations**. In this way, we might be able to explain (and even gain from) possible interactions between the mental and the physical sides of linguistics, in a metatheoretically sound manner. Furthermore, we could postulate that these voluntary operations presuppose the existence of two discrete areas: a number of **inventories** (or 'knowledge bases') where information is to be stored and easily (/quickly) accessible, and a set of **rules of action** handling information coming from and going to these inventories.

The above model has at least one advantage over the 'linguistic' one: it is dynamic, in the sense that voluntary operations over (partly linguistics oriented) inventories accept incessant feedback from and are, therefore, enhanced by the system's general performance (one could compare this model with a -connectionist- parallel distributed network.) What is even more important for its metatheory, this very performance is, strictly speaking, a physical one. There is not enough space here to explain the inner structure of the cognitive phonetic module, but its essence should be clear: it integrates physical (be it vocal-tract or environmental/extra-linguistic) information with the so-called 'central nervous systems' (CNS).

This physical-to-cognitive (and vice versa) interface is best understood if we take account of three important factors:

First, the **relationship between production and perception**; we saw above how phonology evades the (metatheoretical, I argued) problem of relating phonemic and phonetic strings and gave some explanations. When one considers that part of the phonetic representation is implemented voluntarily, autonomously, following non-phonemics-based, yet cognitive, instructions, there seems to be little doubt that phonetics is not, in fact, cannot be assigned or triggered automatically. On the contrary, one could opt for continuously independent deep and surface structures (this goes as back as Bloomfield's [1933] 'continuous derivation'), enabling a specific correlation between *activated* encoding (production) and decoding (perception). Experiments (for a review see Sifakis 1990) have provided evidence for the necessity for the encoding system (operated by the speaker) to know the decoding system's (triggered in the listener) present abilities as experienced by the general speech environment and, secondly, the (pragmatic) means by which such knowledge is to be acquired. Operations carried out by production and perception mechanisms are **compatible** (i.e. they are based on similar, if not identical data stored in the inventories), whereas the encoding and decoding processes themselves are **complementary** (i.e. they are not entirely independent of each other-and see below).

Secondly, the notion of **control**; production and perception systems (i.e. speakers and listeners) must be able to control their operations in a dual way: constructing, on one hand, a long-term method of articulatory control, and achieving, on the other, an on-going acceptable performance. In this way, that is, by continuous trial-and-error tests, both systems collect and store data, thereby enabling quick adaptations to valuable information gleaned from novel 'environmental conditions'. This control is, in the first place, actively operated in production via certain production instructions, processes responsible for the action theory equivalents of 'fine tuning' and 'gross

PHONOLOGY AND PHONETICS – A METATHEORY

control' of the vocal tract musculature (see above for references); in the second place, it is evident in the perception of the speech signal via the ability of the listener to recover the linguistic message beyond phonetic coarticulation and acoustic variability and to adapt to different environments. Needless to say, control is a *learning process* (i.e. human beings are not born fluent speakers of their native tongue -it takes some time before the child comes to terms with his/her physical apparatus which is responsible for speech).

Thirdly, the notion of **choice**; this is probably the most important aspect of cognitive phonetics, since it presupposes the exclusion of phonology (and, hence, of 'linguistic' encoding in the generative sense) from the voluntary act of selecting from within the physically limited set of phonetically (i.e. articulatorily and acoustically) realisable segments those extrinsic allophonic strings that are appropriate to factors such as the speech environment (figures like the amount of familiarity between speaker and listener, whether the speaker has a pipe or experimenting electrodes in his/her mouth, etc., would have to be considered here), the available prosodics and the rate of speech. One should distinguish here between *phonological choice*, which involves a selection of abstract, phonetically constrained yet linguistically motivated objects; *cognitive phonetic choice*, which presupposes the assigning of still abstract, but not linguistics-based (hence, non-phonological) objects based on continuous information flowing in from the environment, the prosodics, etc.; and *phonetic execution*, whose function is that of manipulating the inevitable physical constraints with respect to the cognitive phonetic instructions. The first two are characterised as cognitive (the former being linguistic, the latter the result of voluntary operations), the third as completely automatic.

It may be seen, from the above, that our cognitive phonetic metatheory is in fact capable of bridging the metatheoretical phonology - phonetics gap. This is done by constructing a robust system with its compatible and complementary production and perception sides providing the appropriate input to the automatic phonetics, something that plain phonology cannot do by nature.

4. GENERAL CONSTRAINTS FOR A 'GAP'-ORIENTED METATHEORY

A 'gap'-oriented metatheory should be able to:

(i) give an explanatory account of the way in which human beings recover the linguistic message despite the inherent acoustic variability of physical speech (an account of **perception**);

(ii) give an explanatory account of the way in which human beings encode phonetic variables from phonological constants (an account of **production**);

PHONOLOGY AND PHONETICS − A METATHEORY

(iii) give explicit accounts of how human beings use and/or surpass the problems created by phonetic precision and acoustic variability (cf. Lindblom & MacNeilage 1986);

(iv) account for **phonological acquisition**, that is, the processes (conscious or otherwise) whereby a child comes to know his/her vocal tract constraints and 'creates' the sounds of the language which s/he is exposed to. The metatheory should provide the necessary and sufficient elements for understanding the (historic or otherwise) processes accountable for universal phonetic inventories (in the sense of Lindblom 1991);

(v) provide explanations for the above not only from the **physiology** of speech (its phonetic side), but, also, from its **psychology** (i.e. the cognitive side) and be ready to merge physiological findings with psychological insights;

(vi) give **mathematically accurate** *descriptions* of all the above, possibly via a rewrite (production) rules notation and/or, perhaps, through the performance of an analogically valid computational system (e.g., a neural network, etc.). This also implies its **computational adequacy** (for example, it should be as simple and precise as possible, containing no redundancies, i.e. features or processes explicitly or implicitly reiterated for no fundamental metatheoretic reason; see Marcus 1980: 241*f*.); and

(vii) (especially important for a 'gap' metatheory) abide by the **projectability** criterion (Goodman 1954); according to this methodological constraint, the projectability of a hypothesis can be inherited from the projectability of its vocabulary, and the projectability of an item of scientific vocabulary is accounted for by the frequency with which the item has been projected in previously successful theories. As a result, the history of past theories to some extent determines a present theory's confirmation. The reason why this constraint is of utmost importance for a 'gap'-oriented metatheory is that theorists should not create large amounts of novel jargon which is very likely to be misunderstood by others. Rather, it is best to use successful terms forwarded by older theories, in our case, possibly merging phonological with phonetic jargon. If the outcome is comprehensive, so much the better for the metatheory. (See also Fodor 1983: 108.)

In summary, I have attempted to delineate a cognitive phonetic metatheory which could be described as a linguistics-oriented dualist interactionist infrastructure whose purpose is to create an explanatory and comprehensive (philosophically tenable) psychological-*cum*-physiological background for bridging the phonology - phonetics 'gap'. A number of phonetics- and phonology-oriented theories have been reviewed and rejected on the basis that they are not integrated within any adequate metatheory which would specifically account for the philosophy, psychology and physiology of both phonology and phonetics.

PHONOLOGY AND PHONETICS – A METATHEORY

Anderson, J.M. & Ewen, C.J. (1987) *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Bloomfield, L. (1933) *Language*. London: George Allen & Unwin.

Browman, C.P. & Goldstein, L. (1987) Tiers in articulatory phonology, with some implications for casual speech. *Haskins Laboratories SRSR*, 92/93, New Haven, Connecticut, 1-30.

Caplan, D. (1987) *Neurolinguistics and Linguistic Aphasiology*. Cambridge: Cambridge University Press.

Chomsky, N. (1972) *Language and Mind*. New York: Harcourt Brace Jovanovich.

Chomsky, N. (1980) *Rules and Representations*. Oxford: Basil Blackwell.

Chomsky, N. & Halle, M. (1968) *The Sound Pattern of English*. New York: Harper & Row.

Fodor, J. (1981) The mind - body problem. *Scientific American*, 244, 124-132.

Fodor, J. (1983) *The Modularity of Mind*. Cambridge, Mass.: MIT Press.

Fowler, C.A. (1977) *Timing Control in Speech Production*. Bloomington: Indiana University Press.

Fowler, C.A. (1980) Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 8, 113-133.

Fowler, C.A. (1990) Listener - talker attunements in speech. *Haskins Laboratories SRSR*, 101/102, New Haven, Connecticut, 110-129.

Fowler, C.A., Rubin, R., Remez, R.E. & Turvey, M.T. (1980) Implications for speech production of a general theory of action. in B. Butterworth (Ed.), *Language Production*, vol. 1, London: Academic Press, 373-420.

Fujimura, O. (1991) Comment: Beyond the segment. in I.G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception*, Hillsdale, New Jersey: Lawrence Erlbaum, 27-31.

Giegerich, H. (1985) *Metrical Phonology and Phonological Structure*. Cambridge: Cambridge University Press.

Goldsmith, J. (1979) The aims of autosegmental phonology. in D. Dinnsen (Ed.), *Current Approaches to Phonological Theory*, Bloomington: Indiana University Press, 202-222.

Goodman, N. (1954) *Fact, Fiction and Forecast*. London: Athlone Press.

Hawkins, P. (1984) *Introducing Phonology*. London: Hutchinson University Library.

Hayes, B.P. (1989) *A Metrical Theory of Stress Rules*. Cambridge, Mass.: MIT Press.

Jacobson, R., Fant, G. & Halle, M. (1952) *Preliminaries to Speech Analysis*. Cambridge, Mass.: MIT Press.

Kuhn, T.S. (1970) *The Structure of Scientific Revolutions*. 2nd ed., Chicago: University of Chicago Press.

Ladefoged, P. (1967) *Linguistic Phonetics*. Working Papers in Phonetics, 6, UCLA.

Lass, R. (1984) *Phonology*. Cambridge: Cambridge University Press.

Liberman, A.M., Cooper, F.S., Shankweiler, D. & Studdert-Kennedy, M. (1967) Perception of the speech code. *Psychological Review*, 74, 431-461.

Liberman, A.M. & Mattingly, I.G. (1985) The motor theory of speech perfeption revised. *Cognition*, 21, 1-36.

Lindblom, B. (1991) The status of phonetic gestures. in I.G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception*, Hillsdale, New Jersey: Lawrence Erlbaum, 7-24.

Lindblom, B. & MacNeilage, P. (1986) Action theory: problems and alternative approaches. *Journal of Phonetics*, 14, 117-132.

Lindblom, B., MacNeilage, P. & Studdert-Kennedy, M. (forthcoming) *Evolution of Spoken Language*. Orlando, FL: Academic Press.

Marcus, M. (1980) *A Theory of Syntactic Recognition for Natural Language*. Cambridge, Mass.: MIT Press.

McCarthy, J.J. (1988) Feature geometry and dependency: a review. *Phonetica*, 45, 84-108.

Neisser, U.R. (1967) *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Pavlov, I.P. (1927) *Conditioned Reflexes*. (V. Anrep, transl.), Oxford: Clarendon Press.

Schachter, F.F. (1979) *Everyday Mother Talk to Children*. New York: Academic Press.

Sifakis, N.C. (1990) *A presentation of cognitive phonetics*. Unpublished BA project, Department of Language and Linguistics, Essex University.

Stevens, K.N. (1989) On the quantal nature of speech. *Journal of Phonetics*, 17, 3-45.

Stevens K.N. & Halle M. (1967) Remarks on analysis by synthesis and distinctive features. in W. Wathen-Dunn (Ed.), *Models for Perception of Speech and Visual Form*, MIT Press, Cambridge: Mass., 88-102.

Tatham, M.A.A. (1990) Cognitive phonetics. in W.A. Ainsworth (Ed.), *Speech, Hearing and Language Processing*, vol. 1, London: JAI Press, 193-218.