Variation and linguistic theory

Patrick Honeybone
University of Edinburgh, `patrick.honeybone@ed.ac.uk`

1. Introduction

It may seem surprising, but linguistic variation is often seen as a 'problem' for linguistic theory. The models that formalist, theoretically-minded linguists work with typically assume that linguistic behaviour is categorical and idealise away from the variation that is found in speech. The justification for this, following Chomsky (1965), is that much of the variation found in utterances is due to non-linguistic factors, and thus idealisation is necessary in order to see the underlying patterns behind speakers' linguistic performance. A number of strands of work in theoretical linguistics have, however, sought to take linguistic variation seriously, and they form the topic of this chapter, along with the argumentation that arises when linguistic theorists talk about (or refuse to talk about) linguistic variation.

It's no secret that languages like English are full of variation. If illustration is needed, let us consider a simple sentence like (1), which might describe a woman giving her coat her brother.

(1)   Betty took off her coat and gave him it.

If we limit ourselves to syntactic and phonological variation (as I do throughout this chapter), we could imagine a number of ways in which speakers of English might utter (1), or something very close.[1] As a speaker of English born and raised in the English East Midlands, I could easily utter (1), but I could also utter (2).

(2)   Betty took her coat off and gave him it.

Is (2) the same sentence as (1)? It would be true under the same set of circumstances and it features the same set of words, so let's assume that it is. This means that a speaker from the North-West of England, for example, would also be uttering the same sentence if they said (3), which is how they, among others, might prefer it (as Siewierska & Hollmann 2007 explain).

(3)   Betty took her coat off and gave it him.

The fact that (1), (2) and (3) are all possible in English plainly shows that syntactic variation exists, and variation at the phonological level is also unavoidable. That same speaker from England's North-West might well pronounce *Betty* as [bɛtɪ], *her* as [ə] and *off* as [ɒf]. Some speakers from that area, however, particularly if they came from Lancashire, might pronounce *her* as [əɹ]. Indeed, the same speaker might sometimes pronounce it as [ə] and sometimes as [əɹ]. A speaker from the North-West of the USA, on the other hand, would certainly have a rhotic pronunciation of *her*, like the Lancashire speaker, but would likely pronounce *Betty* as [bɛɾi] and *off* as [ɑf].

Such examples can be multiplied manifold, as any speaker of English knows. *Betty* would likely be [bɛʔi] for a speaker of London English, and many speakers from other parts of the UK might now vary between [bɛti] and [bɛʔi], with different types of populations favouring either the oral stop over the glottal or vice versa. Speakers from Liverpool, on the other hand, may realise the /t/ as a slit alveolar fricative, in a case of lenition, which we can represent as in [bɛθ̠i], and speakers from Newcastle upon Tyne might pronounce the name as [bɛtʔi].

It is often said that linguistic variation occurs when one meaning can be attached to more than one form. This is clearly the case for *Betty*: it doesn't change the meaning if a speaker says [bɛti] one minute and [bɛʔi] the next, and it also seems right to say that 'gave him it' and 'gave it him' *mean* the same, but involve different linguistic forms. These are two cases of linguistic variables – single linguistic items ('meanings') which have multiple identifiable variants ('forms'). The variable (t) has all the variants described above, including [t] and [ʔ] and the variable (pronoun-object-order) has the two variants given here. In this piece, I focus on variation of this type, where one referent has more than one form, and where some sort of geographical, social or at least stylistic effect is associated with the different forms. I leave aside other ways in which language can vary (such as when a phoneme varies categorically in its allophones or a form varies diachronically over time).

How can this abundant variation be a problem for theoretical linguistics? In part, this derives from deep-reaching disagreements about what we mean by 'language'. There is an everyday meaning for that word, and it might seem that that it's obvious what we mean when we talk about 'English'. However, as we will see, neither of these notions are as straightforward as their everyday meanings might imply when we view them through the lens of linguistic theory.

There are two fundamental types of variation that confront us when we consider the notion 'a language like English', and both of them were exemplified above. I investigate this point further in section 2. Sections 3 and 4 take these two types of variation in turn and discuss why they should matter to theoretical linguists, considering some of the methods that are used to analyse such cases of variation, and showing their relevance for linguistic theory. Section 5 concludes.

We cannot all branches of linguistic theory, so I only consider some of the most popular. I use 'linguistic theory' here with its standard, restricted reference: I mean that approach which aims to provide formal, concise statements concerning the structural generalisations that can be made about language in general, or about individual languages. This 'theoretical linguistics' takes at least some impetus from the body of ideas associated with the generative linguistics of Chomsky. It can be contrasted with 'variationist linguistics', in the tradition of Labov (1966), which explicitly focuses on the ways in which speakers vary in their utterances, in terms of the number of variants that they produce for particular linguistic variables. This has shown that all languages are *inherently variable* (including cases of stable variation which can persist in a language for centuries) and that this involves *orderly*

heterogeneity – speakers of similar backgrounds tend to consistently use the same proportion of variants of a variable: variation is not haphazard.

In what is to come, we will both see why much of theoretical linguistics does *not* really pay much heed to linguistic variation (seeing it as a problem which can reasonably be ignored, because the problem actually belongs to someone else – variationists), and consider some work which aims to integrate accounts of linguistic variation into formal linguistic theory (seeing variation as a problem to be solved).

2. Linguistic theory and the two types of variation that it needs to deal with

Language varies in a number of ways, but there are arguably two types of variation which are fundamentally distinct from each other, and which could have different implications for linguistic theory. Some of the variation in (1), (2), (3) and the realisations of *Betty*, *her* and *off*, above, compares forms that are possible in *different* dialects of English, while other aspects of this variation refer to how a *single speaker* (of a single dialect) might realise the forms. The term 'variation' is thus ambiguous, and either (4) or (5) can be intended by it:

(4)   variation between speakers = inter-speaker variation

(5)   variation within a speaker = intra-speaker variation

Inter-speaker variation was illustrated above by the comparison between those speakers who might prefer the order of 'gave it him' and those who might prefer 'gave him it', and by the comparison between those speakers who might tap the /t/ in *Betty* ([bɛɾi]) and those who might glottal it ([bɛʔi]). Intra-speaker variation is involved in the cases where the same speaker might order the particle before the direct object in 'took off her coat' one moment, but might use the other order the next, and in the case where the same speaker might glottal the /t/ in *Betty* in one utterance, but might realise it as a plain [t] in the next.

If our aim is to investigate 'the linguistics of a language' such as English (as it surely is in a volume such as this) we need to consider *both* types of variation as they both exist in the phenomenon that we call 'English'. As we will see when we consider them individually, however – (4) in section 3, and (5) in section 4 – their implications and the responses of linguists who have considered them are very different.

Theoretical linguistics is thus faced with two questions (two problems?): how should it deal with inter-speaker variation and how should it deal with intra-speaker variation? Theoretical linguists of different persuasions have argued that both or neither or only one of these two should be taken into account as they work to figure out the nature of speakers' grammars.

It is probably fair to say that most work in linguistic theory does not see inter-speaker variation as relevant to theory construction. This position depends on answers to two questions, given here in (6) and (7), which go to the heart of a fundamental issue: the nature of the proper object of linguistic study.

(6)   What do we mean by *language*?

(7)   What do we mean by *languages*?

The broadly generative tradition in theoretical linguistics has clear answers to these questions. This tradition includes syntactic work in the narrowly Chomskyan frameworks from the Standard Theory through to Minimalism, and also work in commensurable frameworks like Lexical Functional Grammar and Head-Driven Phrase Structure Grammar. It also includes work in the generative strands of phonology, from Standard Generative Phonology through to Lexical Phonology and Government Phonology, and mainstream work in Optimality Theory (OT) in both phonology and syntax. From within this tradition, Chomsky has long argued that it is crucial to recognise a distinction between different perspectives on (6) and (7). Building on an early distinction between *competence* and *performance*, which is itself reminiscent of Saussure's (1916) distinction between *langue* and *parole*, Chomsky (1986) proposed a distinction between *I-Language* and *E-Language*. I-Language is an aspect of *knowledge* – it is the set of linguistic generalisations that we know in order to be able to speak our language (or languages). For example, we know that in English, the plural is formed by adding an alveolar fricative which takes its voicing from the final segment in the morpheme that it attaches to (unless that segment is a sibilant), and we know that *wh*-phrases like *what* and *how many* occur at the start of clauses in questions. Like all types of knowledge, I-language is *i*-nternal to each *i*-ndividual. It forms the grammar that exists in the mind/brain of speakers. 'E-language' is the term Chomsky uses to refer to all the utterances that speakers make when they talk or write. This connects with everything else that the 'everyday sense' of *language* involves, including those things that are connected to language which are *e*-xternal to the speaker, such as the community that they belong to.

The standard generative opinion is that theoretical linguists should only focus on I-language, and thus the generativists' answer to question (6) is that 'language' only makes sense and can only be a coherently investigable entity if we see it as a cognitive entity that exists in a speaker's mind. E-language is subject to influence from many different areas: physical pressures affecting linguistic performance, social and political pressures affecting interactions, and historical contingency in terms of what speakers happen to talk about. The claim is that the set of utterances that speakers make is not a well-formed or coherent object, and it is not suitable if we hope to discover the generative system that constitutes the grammar. It can't be enough simply to wait for occurrences of everything that is possible in a language in a collection of data: corpora of utterances cannot tell us what is impossible in a language, and so E-language is not a suitable object of study if we hope to figure out the generative system that a speaker knows. On this approach, the answer to question (7) is that 'languages' are I-languages, the grammars which exist in speakers' minds. The implications of this are quite considerable. Each speaker is an individual, of course, and no individuals share a mind. Each speaker's knowledge of language may be slightly different from every other speaker, as their experience in acquisition is different. Thus, while such grammars will need to overlap considerably in order for speakers to communicate with each other, we should not expect them to be the same. Each person has their own language.

All of this shows three things: (i) generative linguistics takes an inherently mentalist approach to language, situating it in the mind; hence, (ii) given that such linguists are required to focus on the knowledge in individual speaker's minds, it cannot make sense to focus on inter-speaker variation; and (iii) the everyday notion of what 'a language' like English is, is seen as incoherent. Points (ii) and (iii) follow from the basic point (i). Regarding point (ii), Chomsky (1965) has famously written that "[l]inguistic theory is concerned primarily with an ideal speaker-listener in a

completely homogenous speech community". This deals with the 'problem' of (4) by identifying it as irrelevant to theory building, so that it should be ignored. The homogeneity that is idealised in order to focus on an individual's grammar might even be taken to exclude intra-speaker variation, but as we will see below, this is not necessary. To consider point (iii), if *languages* are I-languages, which are states of knowledge, there is no room for entities which correspond to the everyday notion of 'English' – something which is shared by millions of speakers, with both small and major variations between speakers or between groups of speakers who speak different English dialects. On this approach "[i]t is very doubtful that one can give any clear or useful meaning to the 'everyday sense' of the term 'language'" (Chomsky, 1980), and the same is true of the term 'dialect'. English is an E-linguistic notion, and there is no point in doing 'the linguistics of a language', because languages do not exist. The only true linguistic object, on this approach, is the idiolect – each speaker's own grammar.

All of these points are challengeable, however, and the answers given above have been rejected by other linguists.[2] While mentalism is the mainstream in theoretical linguistics, other approaches have situated language outside of the mind by giving different answers to questions (6) and (7). An *instrumentalist* approach might argue that language does not exist in the mind because linguists' abstractions and constructs do not (need to) literally exist – theoretical linguistics should simply produce the best analysis of the data that they are presented with and should not care where their generalisations exist, as long as they succeed in characterising the data in the most insightful way and allowing for prediction. Few linguists have overtly adopted instrumentalism, but 'post-Bloomfieldians' in the US, and those following Firth in the UK have done so (see, for example, Carr 2006 and Honeybone 2005). However, some work on linguistic theory can only *seem* to claim a commitment to mentalism, by situating itself in a generative tradition, while the actual analyses proposed seem to hint at a covert instrumentalism. This issue is discussed further briefly in section 3.

More importantly, variationist linguistics has issues with the I/E-language distinction, because variation in production – 'E-language' – is the main focus of study. Labov and others who work in this field aim to discover the patterns that exist in the use of variants that we find in a whole speech community. On this approach, language – the object of linguistic study – exists in the speech community, rather than simply in the I-linguistic mental grammar of each speaker, and some variationist work therefore rejects the focus on I-language, focusing instead on inter-speaker variation, and answering questions such as: in those communities where /t/ can be glottalled, do females glottal more or less than males, or the young more than the old? There *can* also be a focus on intra-speaker variation in Labovian work, however, as we shall see in section 4. Such work aims to model the variable performance of individual speakers.

In the next section, we shall see that, despite protestations to the contrary, inter-speaker variation can and has affected analyses within linguistic theory. This can rely on an overt commitment to a fundamental approach to the conception of language, or it can rely on an analyst simply aiming to come up with the best analysis of variable data.

3. Inter-speaker variation and linguistic theory

As section 2 has shown, most work in theoretical linguistics sets aside inter-speaker variation, as it is not seen to be able to provide relevant evidence for a linguist in their quest to work out a model of an I-language. So can linguistic theory entirely ignore the potential problem of inter-speaker variation? One strand of work within generative linguistics has claimed, to the contrary, that linguistic theory should seek to accommodate variation of this type within the modelling of speakers' grammars. This is *generative dialectology*, which developed in the 1960s and 70s in such work as Newton (1972), and we consider it here. This section would be quite short if that were the only work to consider here because explicitly identified generative dialectology petered out by the 1980s. We shall see, however, that the fundamental ideas that generative dialectology espoused have been and are adopted in other linguistic work. While they are often rejected out of hand in their full-blown, explicitly-stated version, aspects of them can creep into linguistic analyses, and there are resonances of them in a number of areas of theoretical work on English.

Although we have seen that languages such as 'English' have a problematic status in theoretical linguistics, most speakers have the perception that they typically understand other speakers of the thing that we call 'English', even though their speech patterns show considerable variation. Indeed, it could even be argued that languages exist mentally in speakers' perceptions of their close and extended speech community – who it is that they would at least try to speak to if they met. Given this, it is not unreasonable to wonder how speakers store the necessary knowledge of varieties of English that are different to their own. Is this knowledge stored as part of their own I-language? That might be the most economical option, so it is worth entertaining as an idea, and this is what generative dialectology and related approaches do: the grammars that are proposed in such models are meant to account for more than one dialect of a language, with only minimal changes to the analysis to account for the variation. It will be clear that such approaches explicitly aim to account for inter-speaker variation.

This is a *polylectal* approach: one grammar should be able to generate the surface forms of more than one *lect* (that is, variety, dialect, sociolect, *etc.*) of a language. Some work from this perspective (such as Bailey 1973, 1996) argues that the best analysis should be *panlectal*, that is, it should account for every variety of a language (using the word in its perceptual, everyday sense). Polylectal approaches need to combine the facts that (i) something among all the lects that are to be collected in a grammar must be the same (so that all speakers of all the lects have something in common) with (ii) the obvious fact of variation in utterances. The approach that has typically been adopted, in generative dialectology and in conceptually related approaches such as those of Bailey (1972, 1996) and Agard (1971) relies on the widespread analytical assumption that there is more than one linguistically important level. It is often claimed that there is an underlying and a surface level (or a deep and surface structure, or pre- and post-movement stages of a derivation, or an input and output) to any linguistic item, such as a sentence or phonological string. Linguists who adopt a polylectal approach to linguistic analysis have cohered around the idea that the lects of a language like English may vary considerably at the surface level, but are to be expected to be more similar at the underlying level. This makes sense, as it gives a common grammatical 'core' at the underlying heart of all dialects that are perceived to belong to a single language.

Most of the arguments in this area have been made from phonology, so that will be our focus here. As Harris (1984) shows in some detail, however, the same

arguments can be had in syntax. This kind of argumentation fits easily into rule-based phonology, as there is often expected to be something of a gulf between the underlying representation (UR) and the surface representation (SR). All that a polylectal approach needs to assume is that lects of a language differ in terms of the rules that they have (or perhaps in their ordering). In syntax, this type of analysis fits more clearly into early generative models, where surface forms were directly generated from deep structures, but it can be transferred to any syntactic (or phonological) model with 'earlier' and 'later' stages of a derivation. It fits less easily with Optimality Theoretic models, because the input is less important or firmly fixed than is the UR in rule-based phonology, for instance, but it is possible to imagine a model in which the input is expected to remain identical across lects, and only the ranking of constraints should be expected to differ.

Such polylectal approaches to linguistic theory clearly differ from the 'orthodox' approach to variation sketched out in section 2, which emphasises the separateness of linguistic varieties to the extent that each speaker has their own dialect (indeed, their own language). Because of this, I propose to call the approach which contrasts with polylectalism a *dialectal* approach to linguistic analysis, because each dialect/language is analysed entirely on its own, such that dialects may be expected to vary at the underlying level just as much as at the surface level.

There are thus two *fundamental linguistic approaches* that theorists need to choose between in the face of inter-speaker variation: will they take a dialectal or polylectal approach? In fact, we can differentiate further within these fundamental approaches: both are really clines from extremes to less extreme positions, although there is a conceptual gulf between them. All this is schematised in (8).

(8)

| | panlectal | **polylectal** | ‖ | **dialectal** | idiolectal |

An approach which is polylectal to any extent will write a grammar which in some way makes reference to or aims to account for more than one variety, but this could be only a very minor part of an analysis. Panlectalism, the extreme of this position, will seek to produce one underlying analysis to account for all dialects or varieties of a language, with only superficial surface variation, as we saw above. A dialectal approach will stress a need to ignore all varieties other than the precise one that a linguist is focusing on in any particular analysis. Much work with a commitment of this type does indeed claim to focus on the analysis of one single dialect of a language (in the everyday meanings of those words), analysing the phonology of Liverpool English, for example, as in (Honeybone 2001) or the syntax of Belfast English (as in Henry 1995). Most of the time this is not noticed because the variety in question is 'Standard English' or a reference phonological variety, such as General American (GA) or RP, but it is nonetheless only a step on the way to the extreme of this fundamental approach, which is an *idiolectal* approach. This would take the rhetoric of the I-language-only approach seriously, and focus on the grammar of only one speaker. Placement on one of these clines of fundamental linguistic approach does not have to be an all or nothing affair: it would be perfectly possible to adopt a polylectal approach to one aspect of language's phonology or syntax, for example, and a dialectal approach to another.

Consideration of these kinds of distinctions is not accorded all that much importance in most theoretical linguistics, and while the distinctions are discussed in some work, they are often quickly dismissed. However, it is arguably crucial to bear

them in mind when thinking about the treatment of variation in English (or any other language), because it is otherwise easy to fall into the trap of taking one approach while intending to take another.

We might wonder whether all these types of approach are in any way defensible. Clearly we all have *some* knowledge of lects other than our own. But should this knowledge inform a speaker's grammar, or should we merely assume that it is 'tacked on' to our I-language as an additional aspect of language-related knowledge? On a mentalist perspective, panlectalism seems impossible to defend – a language like English has so many speakers, and so many varieties, that it is absurd to assume that any speaker's grammar could be informed by all varieties. It may be defensible under instrumentalism – if it allows for the best account of the regularities of all the data from English to assume a common underlying analysis for one aspect of its linguistic system, some linguists may be tempted to adopt that analysis (as we will see below when we discuss the English monophthong system). Most theoretical linguists will only entertain mentalism, however. A steady stream of work, such as Harris (1984, 1985), McMahon (1992, 2000) and Lodge (2009), argues against panlectalism, and indeed, for the most part, against even the slightest hint of polylectalism.

Certain analyses that are common in the phonology of English have a somewhat polylectal flavour, however, and it seems likely that a similar situation holds for syntax. Take the English segment which is commonly transcribed as /r/, for example. This is subject to massive inter-speaker variation in SRs, as (9) shows.

(9)
- English English speakers typically use [ɹ]
- North American speakers may use [ɹ] or [ɻ]
- Scottish speakers typically use [ɾ] and [ɹ]
- (some) Northumbrian speakers (still) use [ʁ] (see Påhlsson 1972)
- an increasing number of speakers in (especially South-Eastern) England use [ʋ] (see Foulkes & Docherty 2000)

Nonetheless, all of these are typically assumed to be 'the same thing', by both phonologists and speakers: a realisation of English /r/. The very notion 'English /r/' is a polylectal, if not panlectal concept, and the fricative, tap and approximant articulations by which is it realised have little in common phonetically. This variation does not typically cause problems with comprehension between dialects. It could be argued that a polylectal analysis of /r/ in English makes most linguistic sense – all varieties have underlying /r/, which is realised differently in different dialects. This also shows that the URs assumed in a polylectal approach do not need to be taken from any one dialect: [r] is, of course, an alveolar trill, which is almost only ever used as an uncommon stylistic variant in English. Nonetheless, /r/ is widely seen as the formalisation of the underlying segment, with realisation rules required to map it onto the forms shown in (9). This type of situation is not so unusual. For example, although he is generally anti-polylectal, Lodge (2009) perhaps unintentionally assumes a similar situation for the 'English low front lax vowel' thus: "... /æ/, realized as [æ] (RP), [a] (Manchester) or [ɑ] (Belfast) can be treated as the same vowel ...."

Furthermore, the standard rule-based analysis of the kind of inter-speaker variation considered in section 1 for the realisation of /t/ in *Betty* fits in well with a generative dialectological polylectal approach, as shown in (10), where the dialects involved are the same at the underlying level and only differ at the surface, due to the application of different processes.

(10)

	GA	London	Liverpool
underlying	/bɛti/	/bɛti/	/bɛti/
tapping	[bɛɾi]	-------	-------
glottalling	-------	[bɛʔi]	-------
lenition	-------	-------	[bɛθ̠i]
surface	[bɛɾi]	[bɛʔi]	[bɛθ̠i]

Other phonological models might implement this without rules, but could still retain a polylectal applicability (thus the essence of (10) is assumed in Harris 1994's Government Phonology framework). This analysis also coincides with that of Bailey (1980), who uses precisely this type of data to argue for a panlectal approach. This is perhaps a coincidence of analysis, but it might also indicate a general acceptance, or at least, lack of surprise, at the idea that the UR for a linguistic item should be the same across dialects, in the face of surface variation. It's surely true that we expect to find /t/ in every variety of English, just as we expect to find full series of stops, tense and lax vowels, modal verbs and *wh*-fronting, so when we encounter something that seems to correspond to any of these in a dialect which has not been well analysed, we rush to transfer these categories across to the new variety that is under analysis. This is a polylectal approach to analysis. Knowles (1973) bemoans this in his analysis of Liverpool English, and argues, on a truly dialectal approach (but likely incorrectly) that there are no underlying oral stops in that variety, but rather a different category corresponding to the obstruent in *Betty* on the basis of the fact that there is massive lenition in the variety, affecting segments at all places of articulation, thus the fricative in *Betty* is typically found alongside fricatives in the final and medial obstruents in words such as *back*, *dock, tiger, lad, leader*.

(10) shows that an analysis which would be expected on a polylectal approach can appear far from absurd. Other such analyses are less easy to accept. What about the variation in *off* and *her* discussed in section 1? A GA speaker might say [ɑf] and [hɜɹ], while a London English speaker might say [ɒf] and [ɜ], and a speaker of Scottish Standard English might say [ɔf] and [hɛɾ]. The analysis to derive all of these forms from common underliers is much more complicated. /h/ and /r/ would need to be deleted from the London English form, and either [ɜ] derived from /ɛ/ or [ɛ] from /ɜ/ (all this can be done in powerful phonological models, of course). The vowel in *off* is even less straightforward. Because GA lacks a vowel contrast that is present in LE, certain occurrences of GA [ɑ] correspond with LE [ɒ], as in *off*, *plot* while others correspond with LE [ɑ], as in *father, smart*. This distribution is not predictable so *plot* and *father* cannot both have polylectal /ɑ/ as there would be no way of deriving the LE forms by rule. The contrast would need to be stored in the URs of all varieties, and merged on the surface by an absolute neutralisation rule in GA. The same issue would arise for all cases where any variety of English has a contrast that other varieties lack, meaning that all contrasts in any form of English, even very dialectologically restricted ones such as /u/ : /ɪu/ in Welsh English (as in *through* [θɹuː] vs *threw* [θɹɪu]) would need to be stored in URs for all speakers and removed on the surface by neutralisation rules. This is less compelling as an analysis than was (10), and this is tied to the fact that (10) is a quite concrete analysis, with little distance between URs

and SRs, and with a small number of 'natural' processes in the derivation. Typically, the polylectal approach involves an increase in the abstractness of the URs that are needed, and a greater difference between URs and SRs, because the URs have to be mappable onto the SRs of several varieties. Dialectal approaches tend to be phonologically more concrete (and thus more appealing for those who view abstractness in analysis as a problem).

It may be that different types of dialect differences seem reasonable to different degrees under a polylectal analysis. One influential typology of dialect differences is that discussed in Wells (1982) (following others, including Trubetzkoy 1931). A major distinction is made here between *systemic* and *realisational* differences between varieties. Systemic differences occur when dialects differ in terms of the number of contrasting underlying segments that they have, so that there is a difference in the segmental systems of the dialects. Realisational differences occur when dialects have the same type of contrasting underlying segments, but realise them differently, so that there is a difference in terms of the number or nature of the phonological processes in the dialects. Those dialect differences which submit most reasonably to a polylectal analysis are realisational differences. This is the case with (10), where flapping, glottalling and Liverpool lenition are realisational differences. Systemic differences make much less sense under polylectalism: should we really assume that speakers have a large number of underlying contrasts which never surface? A polylectal account of systemic differences, such as English English /u/ : /ʊ/ *vs* Scottish Standard English (SSE) /u/ (as in *pool* and *pull*), and Welsh English /u/ : /ɪu/ *vs* English English /u/ (as in *through* and *threw*) assumes that absolute neutralisation is a common phenomenon, and that we should allow it in one grammar on the basis of argumentation from a different grammar. This suffers from all the criticism levelled at absolute neutralisation (see, for example, Kiparsky 1968, Hooper 1976, Carr 1993, McMahon 2000), and given that phonology is increasingly surface-orientated (as in OT, for example, see Prince & Smolensky 1993/2004), it will persuade few.

The very notion that we should compare dialects (and/or accents) of one language, and that it is worth distinguishing typologies of dialect differences is a polylectal notion. It is common in work on the phonology of English, however – why compare lects of English, rather that varieties of English and varieties of German, if not because we assume that it will help us understand the structure of any one variety by *seeing it as* a variety of English. Wells (1982) is a towering example of this which has been immensely influential in work on phonological variation in English, especially in Europe. Wells developed a basic framework for discussing the vowels of accents of English (consonants are dealt with somewhat differently) which relies on *standard lexical sets*, referred to by keywords, as in (11), which includes all the sets for the vowels that can occur in stressed syllables, in the order that Wells gives them (he numbers them, so that KIT = 1, DRESS = 2 *etc*).

(11) KIT, DRESS, TRAP, LOT, STRUT, FOOT, BATH, CLOTH, NURSE, FLEECE, FACE, PALM, THOUGHT, GOAT, GOOSE, PRICE, CHOICE, MOUTH, NEAR, SQUARE, START, NORTH, FORCE, CURE

The keywords stand for "(i) any or all of the words belonging to the standard lexical set in question; and (ii) the vowel sound used for the standard lexical set in the accent under discussion" (Wells 1982, 124). While they are intended to be supplementable by other sets if other contrasts are needed for a variety, this set of words can also be seen (and can sometimes be used in analysis) as a panlectal analysis

of the English stressed vowel system – that is, the set of contrasts of which we expect to find at least a subset in any lect of English. Forms of the language involved are expected to fit into these predefined slots. Indeed, if we set out those sets which typically involve lax vowels (the first 6 in Wells' list), as in (12a), it looks very much like the symmetrical set of lax vowels in the basic English vowel system in Giegerich (1992), given in (12b).

(12) (a)   (b)
     KIT   FOOT   ɪ   ʊ

     DRESS   STRUT   ɛ   ʌ

     TRAP   LOT   a   ɒ

In some sense, both of these systems provide an analytical framework which is English-specific, and which is applicable, with a small set of moderations (which are themselves built in to the system) to any and all varieties of English.[3] The analysis of non-low tense vowels in English is another area where polylectal argumentation has played a role in well-known analyses. This is shown in (13), which includes the analysis from Trager & Smith (1951), which has been very influential in America. This proposes an 'overall pattern' for the phonology of English, into which multiple varieties should fit. All three systems in (13) expect all or most forms of English to have contrasts which will fit into their slots.

(13)   Giegerich (1992)   Trager & Smith (1951)   Wells (1982)
       i   u   iy   uw   FLEECE   GOOSE
       e   o   ey   ow   FACE   GOAT

The slots assumed in all three systems allow for considerable surface variation. Thus, to take the high back part of the vowel space, it is no problem to reconcile the fact that RP might have [uː] in words such as *goose, loot*, whereas SSE might have [ʉ]. This can be implemented by realisation rules: for RP, all tense vowels are long, and for SSE, the high back tense vowel is fronted, and the Scottish Vowel Length Rule (see Aitkin 1981, McMahon 2000) will account for its length. Again, this analysis fits with the polylectal approach rather well, and it would fit fine in the clothes of generative dialectology. It succeeds in showing that RP [uː] and SSE [ʉ] count phonologically as the same thing (the high, back tense vowel), and that we should expect speakers of the different dialects to be able to recognise the forms as realisations of what is phonologically functioning as the same thing. It may be that an analysis of this type imports a covert instrumentalism into linguistics: it is the best analysis to fit in with an ideal phonological form, which is symmetrical and economical. These are considerations which make for the neatest analysis, but we cannot be sure that they reflect the cognitive reality that mentalism aims for.

The discussion here has focused on phonology, but similar arguments could be imagined for syntax, as Harris (1984) explores. Klima (1964) gives an early generative syntactic analysis (from the period when phonology and syntax were modelled in essentially the same way, on the basis of linguistic rules) which explicitly adopts polylectalism, dealing with a number of cases of variation including the differences in case-marking between varieties which allow (14) and those which allow (14). The difference is accounted for through a basic grammar, which generates

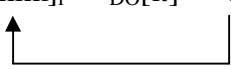(14), and which is supplemented by additional 'extension' rules in lects which allow for (15).

(14) He and I left.

(15) Him and me left.

We could also imagine a movement analysis for object order in ditransitives, which assumes an underlying or initial order such as that in (16), with the direct object closer to the verb, which can surface unchanged as in (3), and also the derivation of (17) in some dialects through movement (leaving a trace), deriving the surface spelt-out form as in (1) and (2), producing a polylectal analysis which shares the approach of those seen above for phonology, above (indeed, Haddican (2010) proposes just such an analysis).

(16) $_{VP}[_V[give]$ $_{DO}[it]$ $_{IO}[him]]$

(17) $_{VP}[_V[give]$ $_{IO}[him]_i$ $_{DO}[it]$ t_i $]$

It is thus certainly possibly to conceive of polylectal analyses in syntax, but little work which overtly identifies with this position exists, in fact. A similar methodological approach to that identified for phonology above can certainly be expected, however, whereby an analyst simply assumes that the categories established on the basis of one set of varieties will be relevant in unresearched varieties of a language.

The above discussion shows that a broadly polylectal approach has been adopted in some work in theoretical linguistics, incorporating knowledge of more than one variety of a language into a single I-linguistic grammar, despite the argumentation in section 2. As we have seen, this has occurred both consciously and sometimes without intention, so that we need to be sure to interrogate an analysis to work out if there are polylectal assumptions behind it.[4]

To sum up this section: how should theoretical linguistics deal with inter-speaker variation? Is a polylectal approach possible? Or should we embrace dialectalism? Or, indeed, idiolectalism? Should linguists allow data and evidence from one variety of a language to weigh on the analysis of another variety in any way at all?[5] Consistent polylectalism (let alone panlectalism) seems problematic, and is widely derided in most work which considers it. There are a number of ways in which arguments stack up against the idea that one I-linguistic grammar can be fully shared across individuals. However, grammatical analysis need not take on or reject every aspect of polylectalism (or dialectalism) wholesale. As we have seen, several widely accepted analyses tend towards polylectal argumentation to some degree. It might yet prove to be a theoretically coherent position to argue for potential partial polylectalism in phonology, with realisational differences between accents candidates for a polylectal analysis, and systemic differences analysed dialectally.

Although it can have an impact on which particular analysis is seen as best for a particular linguistic phenomenon, only little discussion exists of analysts' fundamental linguistic approach to the polylectal/dialectal distinction and the question

of how to deal with inter-speaker variation. The role of intra-speaker variation in linguistic theoretical is typically seen as more important. We turn to this now.

4. Intra-speaker variation and linguistic theory

Like inter-speaker variation, intra-speaker variation in production is a fact of linguistic life. How should linguistic theory deal with it? Some theoretical linguists insist that grammar should only account for what's *possible* in language (such as a glottal stop or plain [t] as a realisation of /t/) and that something else must account for any variation that occurs in the realisation of a category (Hale & Reiss 2008, 139, for example, call this a 'post-grammatical processor'). This would mean that that grammatical theory *should not* account for intra-speaker variation, and that it can be excluded from the argumentation used to work out an I-linguistic grammar. Most theorists who investigate such variation in detail disagree, however, arguing that the same theories that account for the categorical aspects of phonology or syntactic structure should also account for cases where there is intra-speaker variation.

Complex issues can arise when intra-speaker variation and linguistic theory are confronted. As among others Henry (2002) explains, intra-speaker variation has a very uncertain place in the Minimalist Program, the most recent incarnation of Chomskian generative syntax: "Chomsky (1995) proposed an economy principle under which movement occurs only when it is forced to do so. Such a grammar explicitly excludes optionality. If something moves only if forced, it will be impossible in principle for there to be an internalized grammar in which any movement operation is optional. If the option not to move exists, movement will not take place, since it will not be forced to do so." This means that, on a purely formal perspective, intra-speaker variation makes no sense. Minimalism assumes that language is economical, and that its design features should be as minimal as possible for a system that links sound and meaning. "The question arises as to why such a system should have variation – this simply seems to add complications, both in terms of the syntax itself, and of the learnability of the syntax" (Henry 2002).

Speakers do vary in their performance, however. A single speaker may produce (1) or (2), or even (1), (2) and (3), and a single speaker might call the same person both [bɛti] and [bɛʔi] in different utterances. Such variation is crucial socially – it helps people mark out their identity, indicate their sense of belonging to particular places and social groups, and to accommodate to people who they want to identify with, but these are not structural linguistic issues, so they are not easily integrated with the purely formal entities that theoretical linguistics typically deals with, such as segments, syllables, syntactic phrases and functional heads, and some would argue that they absolutely should not be integrated in this way because of this ontological difference.

In what follows we shall (i) briefly consider the observations that have been made about intra-speaker variation, (ii) consider whether such observations should be integrated into formal linguistic analysis and be modelled in the same way as categorical linguistic phenomena, and then (iii) see how this has been done by those who answer (ii) in the affirmative.

Variationist linguistics has shown that, alongside inter-speaker variation, intra-speaker variation is normal in individuals' linguistic performance. As one example among many, Sankoff (2004) shows that one 35-year old male speaker (originally from Yorkshire, but who had moved to the U.S. at age 26) realised words from the STRUT lexical set, which might be described as the variable (ʌ), in the way shown in (18) in a single passage of speech. This speaker clearly realises one category variably.

(18)   ʊ   ɤ   ʌ   phonetic realisation
       4   16   14   number of words

Quantification of the number of tokens of each variant like this is central to variationist sociolinguistics, and it can allow us to measure intra-speaker variation precisely. Such work usually involves more than one speaker, however, producing results which amalgamate the scores of speakers. We can see this in an example of variation at the syntactic level, taken from work by Smith (2000). She discusses variation in the agreement of past tense forms of *be* in Buckie, in North-Eastern Scotland, the variable here being (was/were). Some of Smith's results, as presented in Hudson (2007), are shown in (19).

(19)

Speaker age	You (singular)		We	
	You was	*You were*	*We was*	*We were*
Old	45	5	113	36
Middle	23	12	32	41
Young	43	33	101	45
All	111 (= 69%)	50	246 (= 67%)	122

This shows that, within one age group of speakers, both *you was* and *you were* are found, as are *we was* and *we were*. The numbers in (19) show the total number of times individual forms occurred in the corpus of data that Smith collected, grouped together in three age groups. In ascending age order, the groups contained 9, 14 and 16 speakers, with (nearly) equal numbers of males and females in each group. Quantitative investigations into variation typically present results for groups of speakers in this way, using categories such as age, gender and class (see Guy, this volume). Smith presents her data in a more fine-grained way than (19), but it has been discussed in terms of group scores in the theoretical and variationst literature on variation (for example, Adger 2006, 2007 and Hudson 2007).

What should linguistic theory make of such variation? Although groups of speakers are involved, the approach *need* not be polylectal. Although some researchers in this field do aim to present a picture of the variation found in a community of individuals, others do not. As we saw above, some research explicitly counts the tokens of individuals as well as groups. Other work assumes explicitly that the group scores of quantitative sociolinguistics can be taken to represent the output of individual I-languages. Guy (2005) writes that "extrapolating within communities from group data to individual grammars *is* justified. Granted, it's an approximation, but no more so than any research based on less than perfect sampling." This type of approach aims to propose I-linguistic analyses of variable data, but it may veer into polylectalism as it does so. For the rest of this section, I assume that the theorists are aiming to model I-language when they grapple with variable data, and that they assume that individuals' I-languages *are* involved in the generation of variation.

One position in the theoretical literature which aims to deal with variation with little extra machinery is the claim that all cases of intra-speaker variation occur when a speaker has more than one grammar in their mind. This approach assumes that one grammar produces one variant, such as *you was*¸ or realising /t/ as [t], and this competes with one or more other grammars, which produces the other variants, such

as *you were*, or realising /t/ as [θ̠]. Variation occurs on this model because a speaker sometimes uses one grammar and sometimes another. This approach is found in work such as Kroch (1989), and is described by Lightfoot (1999, 94), discussing variation between OV and VO order in the history of English, which was subject to intra-speaker variation as some points in recorded Old English: "[w]here a language has such an alternation, we say that this manifests diglossia, and that speakers have access to two grammars. Certain speakers have access only to one grammar; other have access to the other grammar; and others have access to both grammars in an internalised diglossia."

At its simplest, this approach means that intra-speaker variation need also not relevant for the analysis of individual grammars – if it is always simply *grammar competition*, each individual grammar can be analysed as an invariable system, fitting in with the impetus within Minimalism to banish grammatical optionality. This position means that intra-speaker variation is actually the same thing as inter-speaker variation on an idiolectal approach: the variation involved simply occurs as the result of different grammars (simply within one speaker, rather than within more than one speaker). For there to be a real conceptual difference between the subjects of this section and of section 3, we can now see that we should not really oppose (4) and (5), but, rather, (20) and (21), where (20) can occur either within or between speakers. (21) is the contentious issue: a model which assumes that intra-speaker variation is explained only by grammar competition, or a 'post-grammatical processor' will claim that (21) does not exist.

(20) variation between systems = inter-grammar variation

(21) variation within a system = intra-grammar variation

It is uncontroversial that a single speaker can possess more than one grammar – this is what we can assume occurs in cases of multilingualism. But is variation of the type shown in (18) and (19) really the same kind of thing as multilingualism? There may well be cases of intra-speaker variation that involve grammar competition, but this seems difficult to defend in many cases of variation. The patterns that are found in Liverpool lenition, for example (as mentioned above), seem to be clearly one system. The stops can be realised (i) as stops in all phonological environments, and can be realised as (ii) stops or affricates in a restricted set of environments, such as a post-nasal context, and can be realised as (iii) stops, affricates or fricatives in a further set of contexts, such as a post-vocalic context (see Honeybone 2001, Watson 2007). This is a classic lenition trajectory (see Honeybone 2008), as shown for /t/ in (22), realising two 'stages' of one lenition process, implying that occurrences of /t/ in the right environment (eg, word-finally in a word such as *dock*) can be realised as [t], [t̠θ̠] or [θ̠].

(22)   t → t̠θ̠ → θ̠

The realisations as affricates and fricatives seem so closely related, with similar but slightly different constraints on their realisation, that it is difficult to believe that [t̠θ̠] is derived by a completely different grammar to [θ̠]. This relevance of this approach is reinforced by cases where the same structural phenomenon can be seen to play both a categorical role and variable role in the same language. Guy & Boberg (1997) argue, for example, that the Obligatory Contour Principle (the 'OCP', which

forbids identical adjacent phonological elements, or forbids them from sharing specific phonological property, such as a particular constellation of phonological features) plays a role in both categorical and variable phenomena in English phonology. Categorically, it forbids lexical geminates and enforces the avoidance of sequences of similar segments in past tense and plural suffixation (with *-ed* and *-es*). They argue that it also plays a role in determining when Coronal Stop Deletion (CSD) is likely to occur in a number of American varieties of English. CSD is one of the best-studied variable phonological processes in English. It involves the deletion of [t] and [d] when they occur in word-final consonant clusters, so that words like *rift*, *west*, *bold*, and *find* can be pronounced with final [f] or [ft], [s] or [st], [l] or [ld], or [n] or [nd], respectively. Intensive study has shown that the likelihood of CSD applying varies according to a number of factors, including the nature of the previous segment (as shown in (23), for Philadelphia English, from Guy & Boberg, 1997) and the nature of the following phonological environment (as shown in the percentage figures in (24) for a range of varieties, from Coetzee & Pater, to appear, who also give references for the studies of each variety). Thus, for example, in terms of preceding segments, the stops are most likely to be deleted if they are preceded by a sibilant fricative or stop (which will be non-coronal, as in *act*), and are least likely to be deleted if they are preceded by a lateral (when they are still likely to be deleted, but could also perfectly well be pronounced).

(23)

Preceding Segment	Probability of Deletion[a]
Sibilant fricatives	1.00
Stops	0.84
Nasals	0.78
Nonsibilant fricatives	0.69
Laterals	0.66
$N = 1,860$	

(24)

	Pre-V west and	Pre-Pause west	Pre-C west side
AAVE (Washington, DC)	29	73	76
Chicano English	45	37	62
Jamaican English	63	71	85
New York City English	66	83	100
Tejano English	25	46	62
Trinidadian English	21	31	81
Philadelphia English	38	12	100

   Guy & Boberg (1997) argue that the generalisations in (23) can be insightfully accounted for if we assume that a probabilistic OCP favours deletion of the following coronal stop if the preceding segment shares the features that characterise the stops (which they assume to be [−son, −cont, +cor]). They further argue that this effect is probabilistic and thus cumulative: the more features that are shared, the more likely it is that the stops delete. Their full analysis is intricate, but a simple form of it will show its basics: the segments most favouring deletion share two features ([−son, +cor] in the case of sibilants, [−son, −cont] in the case of stops), the nasals are mixed in terms of number of features shared, and the segments least favouring deletion share one feature ([−son] in nonsibilant fricatives and [+cor] in laterals). Guy & Boberg argue that the fact that structural linguistic constraints such as OCP can be seen to play a role in both categorical and variable phonology means that I-linguistic grammars should be expected to model categorical and variable phenomena in the

16

same way, with the latter accounting for intra-grammar variation, which is seen as an authentic phenomenon.

A number of approaches have been developed to deal with intra-grammar variation within linguistic theory. The precise implementation of this naturally depends on the characteristics of the theoretical framework, and we cannot consider all linguistic frameworks here, for reasons of space. I narrow the discussion to most popular models: rule-based phonology, Optimality Theoretical phonology and syntax, and principles and parameters syntax (although it should be noted that other models have also been proposed as appropriate to model variability, such as Word Grammar, see Hudson 1997 and HPSG, see Bender 2007).

The simplest way to model variation is to assume that a phenomenon, such as the realisation of /t/ as [ʔ] or [θ], or the deletion of /t/ and /d/ in CSD is controlled by a linguistic rule which is marked as optional in the grammar. The phonological use of optional rules stretches back to Chomsky & Halle (1968), and has recently been defended by Vaux (2008).[6] However, most of the work in rule-based phonology which takes variation seriously has replaced optional rules with variable rules, which aim to write in factors into the formulation of a rule which govern how often particular variants of a variable occur, so that "the predicted relative frequency of a rule's operation is, in effect, an integral part of its structural description" (Cedergren & Sankoff 1974). This is in part based on the claim that variation is not simply a matter of optionality, but is subject to orderly heterogeneity.

Labov (1972) discusses CSD, presenting a number of ever-refined variable rules to account for it. One of these is given in (25), accounting for the observation shown in (24), that the stop is less often deleted if it precedes a vowel than if it does not.

(25)   [−cont] → <∅> / [+cons] __ # # <−syl>

The angled brackets in (25) indicate aspects of the rule that place variable constraints on it, so (25) does not claim that the whole process is optional − rather, it indicates that the process is more likely to apply if a vowel does not follow the word-final stop, producing a probabilistic grammar. The variable rule approach has been taken up in variationist linguistics, but has not been all that widely used, probably because most variationist work focuses on quantifying the variation found in particular linguistic phenomena, rather than modelling it grammatically, and most phonological work does not focus on variation, either dealing only with categorical processes, or setting aside the fact that the processes considered are, in fact, variable. The same could be said for all theoretical approaches considered here, but most of the work on phonological theory which *does* engage with variation is now conducted in OT. Some work on variation and syntactic theory also uses OT, but the principles and parameters framework is also well represented in theoretical discussions of syntactic variation.

OT replaces linguistic rules with violable constraints on linguistic forms which are ranked in terms of their relative order of importance in determining the surface form (called the output). The earliest variable OT work was similar to the optional rule format in simply allowing more than one output without giving any indication of the pattern in the variation, thus an analysis could allow two candidates to tie in terms of constraint violations, either because neither violates a set of fully ranked constraints, or because particular constraints are tied in the ranking. This is illustrated in (26), from Pesetsky (1997), which accounts for variability in the realisation of the declarative complementiser *that*. Both forms of the sentence in (26) are possible, with

the complementiser pronounced (26a) or not (26b), and Pesetsky accounts for this by giving them the same number of constraint violations, of equally ranked constraints (LE(CP) and TEL are 'tied constraints'). The constraints are:

- RECOVERABILITY = a syntactic unit with semantic content must be pronounced unless it has a sufficiently local antecedent
- LEFTEDGE(CP) = the first pronounced word in CP is the complementiser that heads it
- TELEGRAPH = do not pronounce function words

(26)

Candidates	REC	LE(CP)	TEL
a. I believe [cp that Peter is hungry]. [on ranking: LE(CP)>>TEL] ☞			*
b. I believe [cp ~~that~~ Peter is hungry]. [on ranking: TEL>>LE(CP)]☞		*	

This approach has not found widespread favour within OT, however, because there will likely always be some constraint, even if it is very low ranked, which would be able to decide between the candidates which seem to be tied in the necessarily abbreviated constraint rankings that are considered in analyses such as (26). A more finessed approach to variation in OT assumes that constraints can be variably ranked. Such approaches always deliver a winning candidate, avoiding the problem of tied constraint violations. They assume that, while most constraints are ranked normally, the constraints which account for variation are only partially ordered – they 'float' in a block in the hierarchy and are not ranked with reference to each other.[7] Each time that the grammar is used to evaluate a candidate set, the partially-ordered constraints are fully ranked in a way which is consistent with the partial ordering – the ranking is randomly chosen, by the grammar, delivering surface variation.

Unlike the tied-constraint-violation model, this approach aims to say something about the probability with which particular variants occur (thus improving on the 'tied-violation' model, as variable rules aimed to improve on optional rules). It aims to model the type of quantification of variation shown in the numbers given in (18), (19), (23) and (24).

The number of grammars which can be created through the random ranking of a set of partially ordered constraints is limited, and an analyst can work out how many of each of the possible grammars allows for each possible variant output. As the ranking is entirely random, the prediction is that the number of possible grammars which gives a variant output should be align with the proportion of variants actually produced by speakers. Coetzee & Pater (to appear) describe Coetzee's (2004) model (following work by Kiparsky and Anttila) of the aspects of CSD shown in (24). The constraints used are:

- *CT = consonant clusters may not end in a coronal stop
- MAX = input consonants must be present in the output
- MAX-PRE-V = input consonants in pre-vocalic position must be present in the output
- MAX-FINAL = input consonants in phrase final position must be present in the output

If these constraints are variably ranked, they can give rise to 24 different rankings, which fall into five sets in terms of the outputs they can produce, shown in

(27). If MAX is ranked highest, as in (27a), there can be no deletion, no matter how the other constraints are ranked, and if *CT is ranked highest, (27e), the stops will delete no matter whether they are pre-vocalic, pre-consonantal or pre-pausal (phrase final). Other rankings show different patterns of deletability.

(27)

			Deletion produced?		
	Crucial rankings	Total # rankings	Pre-V	Phrase-final	Pre-C
a.	MAX >> *CT	12	No	No	No
b.	MAX-PRE-V >> *CT >> {MAX, MAX-FINAL}	2	No	Yes	Yes
c.	MAX-FINAL >> *CT >> {MAX, MAX-PRE-V}	2	Yes	No	Yes
d.	{MAX-PRE-V, MAX-FINAL} >> *CT >> MAX	2	No	No	Yes
e.	*CT >> {MAX, MAX-PRE-V, MAX-FINAL}	6	Yes	Yes	Yes

The prediction here (shown by comparing the second two columns) is that pre-consonantal deletion is much more likely than deletion in a pre-vocalic or phrase-final environment because many more grammars allow it, and this matches the observed results in (24) for all dialects. However, the precise percentages of the observed variation are not modelled, nor could they be using the same set of constraints for all dialects, as the proportions are different. (27) produces the same number of rankings to allow deletion in pre-vocalic and phrase-final environment. (24) shows that dialects vary as to which of these two environments is most likely to allow deletion, so this may be the right result, but it means that it is not clear how far the model can or should go in modelling the probability of particular variants occurring. It would be too optimistic to expect that such modelling will always be possible to account for the number of variants of each variable, and the variation that is captured here cannot model differences between individuals in a speech community – it only models broad relative probabilities in the overall surface variation, relying on fixed numbers of specific constraints to derive the numerical effects. Problems with this approach are perceivable: if there were also a MAX-PRE-C, the numbers would not fit so well. Also, some work in this approach goes further, attempting to model precise percentages of variation in terms of the likelihood of particular rankings, relying on the existence of precise numbers of constraints for each case of variation.

The other main OT model of variation retains the idea that there is always one full constraint ranking for each 'use' of the grammar but models variation more simply, by essentially weighting constraints with numbers which determine how likely a particular constraint ranking is. Constraints can be allowed to overlap in their ranking (to different, numerically-specified degrees), as in Stochastic OT (eg, Boersma & Hayes 2001), such that one constraint can overpower another most of the time, but not all of the time, in line with their ranking value and some random numerical 'noise', which is generated each time a linguistic form is processed, ready for utterance. This approach is quite radical, as it incorporates numbers directly into the grammar, meaning that language acquirers must abstract the ranking value for constraints from the speech that they are confronted with, and attach it to a particular constraint, and, perhaps for this reason, has met with some considerable rejection.

Stochastic OT was developed in phonology, but has also been applied to syntax, as in Bresnan, Deo & Sharma (2007). Most work in theoretical syntax works with non-violable principles which apply in all languages in the same way, however, rather than

ranked constraints. Cross-linguistic variation is derived largely through the use of parameters, which allow a linguistic system to choose from a restricted range of structural possibilities. When syntax was modelled using construction-specific rules, it is easy to imagine how optional or variable rules could be used in the same way as they were in phonology, as discussed above in (25), and, although little variationist work on syntax was being carried out at that point, obvious cases of variation, such as in verb-particle constructions, like the difference between (1) and (2), could simply be modelled by an optional rule (as could passivisation, for example).

In frameworks which work with principles and parameters, intraspeaker variation has often been modelled as grammar competition involving multiple grammars within one speaker, as in work by Kroch and Lightfoot, discussed above. Not all work has assumed this position, however. Henry (1995), and considerable subsequent work, has argued that the grammar competition model seems implausible when confronted with multiple cases of variation. Thus, Henry (2002) writes that "within Belfast English, there is variability in relation to agreement patterns, word order in imperatives, inversion in embedded questions, and the use of the relative pronoun in subject contact relative clauses. By no means every speaker has variation in all of these, so that there is a range of possible grammars with and without variation for a range of structures; if there is grammar competition, then it is between a wide range of grammars, not just two, and a better characterization seems to be that individual structures/parameter settings are variable, rather than that there are actually separate grammars."

Henry (1995) argues that the approach later adopted by Minimalism – that intra-grammar variation is impossible – is problematic. She argues explicitly that there can be optionality in syntactic movement. For example, certain dialects of Belfast English allow overt subjects in imperatives (a clear inter-speaker difference to Standard English), and they furthermore allow weak object pronouns to precede or follow the subject, thus both (28a) and (28b) are possible as imperatives.

(28) (a)   Give it you to the teacher.
     (b)   Give you it to the teacher.

Henry accounts for this difference as optional subject raising, from SPEC/VP to SPEC/AGR$_S$P. She argues that similar cases of syntactic optionality should be modelled in similar ways, typically assuming that differences of parameter settings drive the variation This leaves the likelihood of either form occurring fully outside of the grammar, drawing back to a position where grammar should only account for what is possible, leaving sociolinguistic factors (a 'post-grammatical processor'?) to account for which form is used when.

Other syntactic work has tried to retain the notion that syntactic structure does not allow for optionality by exploiting the reinterpretation of parameters that has occurred in principles and parameters syntax, so that they are no longer seen as 'switches' but as lexical items. As Adger and Trousdale (2007) write, this "doesn't seem like a major change; however, while it is impossible to have a switch being in both positions at once, it is certainly possible to have two lexical items in a language with contradictory specification." This means that it is possible to model variation such as the variation in agreement described in (19), as in Adger & Smith (2005) and Adger (2006), by allowing a language to contain both agreement options as functional elements, which a speaker tacitly chooses from when a linguistic form is processed.

Adger (2006) also aims to model aspects of the probability of occurrence of particular forms.

This section has shown that linguistic theorists have developed a range of ways of dealing with the problem of inter-speaker variation, even if it is seen as true intra-grammar variation. The question as to whether theoretical linguistics *should* deal with this type of variation remains unanswered in many phonologists and syntacticians minds, however. Linguistic theory has had a clear degree of success in modelling the patterns of variation that can occur, and that have been observed in variationist linguistics, as we have seen. Many would still argue, however, that, while linguistic theory might allow for and model variation, it should not predict how frequent each variant of a variable will occur, because that is seen as purely the job of sociolinguistics. The status of variation in syntax has also led to some concern – it is less easy to see two different surface forms as being derived from the same underlying form in syntactic theory than it is in phonology. It is more straightforward to see a statement that "the variable (t) has the variants [t] and [ʔ]" as being equivalent in status to a rule such as /t/ → [ʔ], but the equivalent does not sit so easily with modern models of syntax. We have assumed in this chapter that essentially the same issues arise when theoretical phonology and syntax are confronted with variation, but some syntacticians would object that this is not so.

Intra-speaker variation is more widely seen as something which linguistic theoretical linguistics should deal with than is inter-speaker variation. The issues that the two raise are rather different. As we have seen, intra-speaker variation only really forces linguists to develop new theoretical machinery if it is accepted that it involves intra-grammar variation. If it is, the models discussed in this section become important extensions of linguistic theory.

5. Conclusion

This chapter has considered quite a wide range of issues which arise when linguistic theory is confronted with the observations that variationist linguistics and other studies of linguistic variation, such as dialectology, have established. Both inter-speaker and intra-speaker variation can be seen to pose problems for linguistic theory, and, as we have seen, some theorists consciously decide to ignore these problems, believing that theoretical linguistics should deal only with categorical phenomena, or at most should describe what is possible, not what is likely or unlikely (for example, Newmeyer 2003, 2005 summarises a range of arguments as to why theoretical linguistics should ignore variation). We have also seen that some considerable thought has been devoted to tackling these problems head-on, however. While linguistic theory wrestles with these issues, and there is substantial disagreement in the field about how linguists of a theoretical bent should deal with them, I hope to have shown that variation is not really a *problem*, but rather an opportunity for innovative thinking, to test the boundaries of linguistic theory and for far-reaching argument concerning the nature of language.

6. Where to go next

For empirical details of phonological variation between speakers (and varieties) in English, Wells (1982) is still the place to start, but this should be read in conjunction with more detailed work, such as the chapters in Foulkes & Docherty (1999) for UK varieties and Labov, Ash & Boberg (2006) for US varieties. Variation in both

phonology and syntax are covered in the chapters in the two-volumed Kortmann & Schneider (2004). There is a vast amount of variationist quantitative work on English, which is relevant to understanding variation within as well as between speakers, starting with that by Labov (Labov 1972, for example, is still well worth reading). Any textbook on sociolinguistics (such as Chambers 2003) will likely also discuss the basics. Most relevant is work applying linguistic theory to data which involves variation in some way. For syntax, the articles in Trousdale & Adger (2007) and Cornips & Corrigan (2005) represent a good cross-section of work, and the introductions to the volumes discuss fundamental issues and details, as does the overview in Henry (2002). For phonology, the chapters in Hinskens, van Hout & Wetzels (1997) are similarly worth consulting, as are overviews like Coetzee & Pater (to appear) and Anttila (2002). There is a considerable strand of work attacking the polylectal approach to inter-speaker variation: parts of Harris (1984, 1985), McMahon (1992, 2000) and Lodge (2009) all consider some aspects of the issues. Little work explicitly advocates polylectalism (the work of Bailey, such as Bailey 1996 is one often-cited exception, but is unfortunately difficult to read and contains few analyses). However, as mentioned in the chapter, traces of this approach can be discerned in quite a range of work.

7. Bibliography

Adger, D. & Smith, J. (2005) Variation and the Minimalist Programme. In L. Cornips & Corrigan, K. (eds) *Syntax and Variation: Reconciling the Biological and the Social*. Amsterdma: John Benjamins. 149-178

Adger, D. (2006) Combinatorial Variation. *Journal of Linguistics* 42. 503-530.

Adger, D. (2007) Variability and Modularity: a Response to Hudson. *Journal of Linguistics* 43. 503-530.

Adger, D. & Trousdale, G. (2007) Variation in English Syntax: Theoretical Implications. *English Language and Linguistics* 11. 261-278.

Agard, F. (I971) Language and dialect: some tentative postulates. *Linguistics* 65. 5-24.

Aitken, A. J. (1981) The Scottish Vowel-Length Rule. In Benskin, M., & Samuels, M. (eds.) *So meny People, Longages and Tonges*. Edinburgh: The Middle English Dialect Project. 131-157.

Anttila, A. (2002) Variation and Phonological Theory. In Chambers, J., Trudgill, P. & Schilling-Estes, N. (eds.) *Handbook of Language Variation and Change*. Oxford: Blackwell.

Bailey, C.-J. (1973) *Variation and Linguistic Theory*. Arlington: Center for Applied Linguistics.

Bailey, C.-J. (1980) Conceptualizing dialects as implicationsl constellations rather than as entities bounded by isoglossic bundles. Chapter 4 of Bailey (1996). Originally in Göschel, J., Ivić, P. & Kehy, K. (eds) *Dialekt und Dialektology*. Wiesbaden: Franz Steiner.

Bailey, C.-J. (1996) *Essays on Time-Based Linguistic Analysis*. Oxford: Oxford University Press.

Bender, E. (2007) Socially meaningful syntactic variation in sign-based grammar. *English Language and Linguistics* 11. 347-381.

Boersma, P. & Hayes, B. (2001) Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32. 45-86.

Bresnan, J., Deo, A. & Sharma, D. (2007) Typology in variation: a probabilistic approach to *be* and *n't* in the Survey of English Dialects. *English Language and Linguistics* 11. 301-346

Carr, P. (1993) *Phonology*. Basingstoke: Macmillan.

Carr, P. (2006) Philosophy of Linguistics. In Brown, K. (ed) *Encyclopedia of Language & Linguistics*. Second edition. Amsterdam: Elsevier.

Cedergren, H. J. & Sankoff, D. (1974) Variable rules: performance as a statistical reflection of competence. *Language* 50. 333-355.

Chambers, J. (2003) *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. Second edition. Oxford: Blackwell.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Chomsky, N. (1986) *Knowledge of Language: its Nature, Origin, and Use*. New York: Praeger

Chomsky, N. (1995) *The Minimalist Program*. Cambridge: MIT Press.

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

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

Coetzee, A. (2004) *What it Means to Be a Loser: Non-Optimal Candidates in Optimality Theory*. PhD thesis, University of Massachusetts Amherst.

Coetzee, A. & Pater. J. (to appear) The place of variation in phonological theory. In Goldsmith, J. Riggle, J. and Yu (eds.) *The Handbook of Phonological Theory*. Oxford: Blackwell.

Cornips, L. & Corrigan, K. (2005) *Syntax and Variation: Reconciling the Biological and the Social*. Amsterdam: John Benjamins.

Foulkes, P. & Docherty, G. (eds.) (1999) *Urban Voices: Accent Studies in the British Isles*. London: Arnold.

Foulkes, P. & Docherty, G. (2000) Another chapter in the story of /r/: 'labiodental' variants in British English. *Journal of Sociolinguistics* 4. 30-59.

Giegerich, H. (1992) *English Phonology: an Introduction*. Cambridge: Cambridge University Press.

Guy, G. & Boberg, C. (1997) Inherent variability and the obligatory contour principle. *Language Variation and Change* 9. 149-164.

Guy, G. (2005) Grammar and usage: a variationist response. *Language* 81. 561-563.

Hale, M. & Reiss, C. (2008) *The Phonological Enterprise*. Oxford: Oxford University Press.

Harris, J. (1984) Syntactic Variation and Dialect Divergence. *Journal of Linguistics* 20. 303-327.

Harris, J. (1985) *Phonological Variation and Change*. Cambridge: Cambridge University Press.

Harris, J. (1994) *English Sound Structure*. Oxford: Blackwell.

Henry, A. (1995) *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford: Oxford University Press.

Henry, A. (2002) Variation and Syntactic Theory. In Chambers, J., Trudgill, P. & Schilling-Estes, N. (eds.) *The Handbook of Language Variation and Change*. Oxford: Blackwell.

Hinskens, F., van Hout, R. & Wetzels, L. (1997) *Variation, Change and Phonological Theory*. Amsterdam: John Benjamins.

Honeybone, P. (2001) Lenition Inhibition in Liverpool English. *English Language and Linguistics* 5. 213-249.

Honeybone, P. (2005) J.R. Firth. In Chapman, S. & Routledge, C. (eds) *Key Thinkers in Linguistics and the Philosophy of Language*. Edinburgh: Edinburgh University Press. 80-86.

Honeybone, P. (2008) Lenition, weakening and consonantal strength: tracing concepts through the history of phonology. In Brandão de Carvalho, J., Scheer, T. & Ségéral, P. (eds) *Lenition and Fortition*. Berlin: Mouton de Gruyter. 9-93.

Hooper, J. (1976) *An Introduction to Natural Generative Phonology*. New York: Academic Press.

Hudson, R. (1997) Inherent variability and linguistic theory. *Cognitive Linguistics* 8. 73-108.

Hudson, R. (2007) English dialect syntax in Word Grammar. *English Language and Linguistics* 11. 383-405.

Kiparsky, P. (1968) *How abstract is phonology?* Bloomington: Indiana University Linguistic Club.

Klima, E. (1964) Relatedness between Grammatical Systems. Language 40. 1-20

Knowles, G. (1973) *Scouse: the Urban Dialect of Liverpool*. PhD thesis, University of Leeds.

Kortmann, B. & Schneider, E. (eds) (2004) *A Handbook of Varieties of English*. 2 volumes. Berlin: Mouton de Gruyter.

Kroch, A. (1989) Reflexes of grammar in patterns of language change. *Language Variation and Change* 1. 199-244.

Labov, W. (1966) The social stratification of English in New York City. Washington DC: Center for Applied Linguistics

Labov, W. (1972) Sociolinguistic patterns. Philadelphia : University of Pennsylvania Press.

Labov, William, Sharon Ash, and Charles Boberg (2006). The Atlas of North American English. Berlin: Mouton de Gruyter

Lightfoot, D. (1999) *The Development of Language: Acquisition, Change, And Evolution*. Oxford: Blackwell.

Lodge, K. (2009) *Fundamental Concepts in Phonology: Sameness And Difference*. Edinburgh: Edinburgh University Press.

McMahon, A. (1992) Underspecification theory and the analysis of dialect differences in lexical phonology. *Transactions of the Philological Society* 90. 81-119.

McMahon, A. (2000) *Lexical Phonology and the History of English*. Cambridge: Cambridge University Press.

Newmeyer, F. (2003) Grammar is grammar and usage is usage. *Language* 79. 682-707.

Newmeyer, F. (2005) A reply to the critiques of 'Grammar is grammar and usage is usage'. *Language* 81. 229-236.

Newton, B. (1972) *The Generative Interpretation of Dialect: a Study of Modern Greek Phonology*. Cambridge: Cambridge University Press.

Påhlsson, C. (1972) *The Northumbrian Burr: a Sociolinguistic Study*. Lund: Gleerup.

Pesetsky, D. (1997) Some Optimality principles of sentence pronunciation. In Barbosa, P., Fox, D., Hagstrom, P., McGinnis, M. & Pesetsky, D. (eds) *Is the Best good Enough? Optimality and Competition in Syntax*. Cambridge: MIT Press.

Prince, A. & Smolensky, P. (1993/2004) *Optimality Theory: Constraint Interaction in* Generative *Grammar*. Rutgers University & University of Colorado at Boulder. Published 2004, Oxford: Blackwell.

Sankoff, G. (2004) Adolescents, young adults and the critical period: two case studies from Seven Up. In Fought, C. (ed) *Sociolinguistic Variation: Critical Reflections*. Oxford: Oxford University Press.

Saussure, F. de (1916) *Cours de linguistique générale*. Edited by C. Bally & A. Sechehaye, with the collaboration of A. Riedlinger. Paris: Payot.

Siewierska, A. & Hollmann, W. (2007) Ditransitive clauses in English with special reference to Lancashire dialect. In Hannay, M. & Steen, G. (eds.) *Structural-Functional Studies in English Grammar*. Amsterdam: John Benjamins. 83-102.

Smith, J. (2000) *Synchrony and Diachrony in the Evolution of English: Evidence from Scotland*. PhD thesis, University of York.

Trager, G. & Smith, H. (1951) *An Outline of English Structure*. Norman: Battenburg.

Trousdale, G. & Adger, D. (eds.) (2007) Special Issue on English Dialect Syntax: Theoretical Perspectives. *English Language and Linguistics* 11(2).

Trubetzkoy, N. (1931) Phonologie et géographie linguistique. *Travaux du Cercle Linguistique de Prague* 4. 228-234.

Vaux, B. (2008) Why the phonological component must be serial and rule-based. In Vaux, B. & Nevins, A. (eds.) *Rules, Constraints, and Phonological Phenomena*. Oxford: Oxford University Press. 20-61.

Watson, K. (2007) *The Phonetics and Phonology of Plosive Lenition in Liverpool English*. PhD thesis, Edge Hill College/Lancaster University.

Wells, J. (1982) *Accents of English*. Three volumes. Cambridge: Cambridge University Press.

Additional Endnotes

[1] The restriction of discussion to phonology and syntax fits in with historical and current trends in the aspects of linguistics that are our focus. A clear characteristic of theoretical linguistics is its syntactocentrism − that is, the primacy of syntax. This can be understood in a number of ways: it is common in generative approaches to see syntax as the one linguistically essential grammatical system, which feeds forward syntactically-derived structures to phonology and semantics, with the latter two being seen as (only) interpretive systems (see Jackendoff 2002, for example). Perhaps tied to this, the volume of work in theoretical syntax dwarfs that in other areas of theoretical linguistics. Opposing this, there is a clear phonocentrism in variationist linguistics (and also in work which positively aims to engage with both variation and formal linguistic theory). Most ideas about how linguists should deal with variation were first developed in phonological work (often dealing with variation in the phonology of English, see Carr & Honeybone 2007). This work properly started in the 1960s, and was joined only later by work on variation in syntax (such as the variationist Cheshire 1982, and the formalist Henry 1995), and this bias towards phonological work in discussion of variation is unavoidably reflected in the chapter below.

[2] One partial challenge is that it can be argued that languages exist, and that they exist *mentally*, on the grounds of speakers' perceptions. People have conceptions of who speaks the same language as them − that is, of who belongs to their close or extended speech community. These are the speakers who they would try to understand when they meet (even if is sometimes proves very difficult when their dialects are very different). The same person will not normally expect to be able to, or try to, understand a speaker of what they perceive to be a different language, even if it is a language that is historical or synchronically similar to their own. This all seems to imply that speakers *do* have mentally real conceptions of 'languages' like English. It might be that it is these perceptions (which can be studied through the methodology of perceptual dialectology, see Preston 1989) that we can point to, to say that this is where languages (in the everyday sense) exist. This knowledge of languages may be of a very different kind to a Chomskyan knowledge of an I-linguistic structural generative system (a grammar), but they both exist in the same mind/brain. It may yet be that one can influence the other.

[3] Thus the Northern English English absence of contrast between the vowels in *put* and *putt* is describable under (12a) by using the same vowel in the FOOT and STRUT sets, or under (12b) by the absence of the /ʌ/ vowel, but the slots for potential variation of this sort are part of the analyses. Wells avoids having to use particular phonological symbols in this, which can be problematic as it could be understood to imply that the variety whose vowels are closest to the symbol used is the 'best' or most fundamental (and this is what Lodge does when he says that RP-based /æ/ is realised as [a] in Manchester), but this is not unavoidable if symbols are used for the slots, as long as they are not simply taken from one variety, as in the example in (13).

[4] Wolfram (1982) has shown that speakers *can* have knowledge of linguistic phenomena which are not part of their variety of a language, by investigating the intuitions of middle class residents of Washington, D.C. about Appalachian English *a*-prefixing of *-ing* forms. Wolfram shows that there are a number of subtle constraints on *a*-prefixing which are part of the grammar of speakers of Appalachian English. For example, it is possible in progressives, e.g. (1), and adverbial complements (2), but not in gerund constructions following a preposition (3), although corresponding structures without an overt preposition are possible (4). Phonological constraints also apply in *a*-prefixing: it is not possible in (5), when the first syllable of the root is unstressed, for example.

(1)   I knew he was a-tellin' the truth, but still I was a-comin' home.
(2)   You was pretty weak by the tenth day, a-layin' there in bed.
(3)   *He got sick from a-workin' so hard.
(4)   He got sick a-workin' so hard.
(5)   *He was a-returnin' from the house.

Wolfram (1982) shows that the speakers of English from Washington, D.C., who do not have *a*-prefixing in their production, performed in almost exactly the same way as native Appalachian English speakers in a grammaticality judgement test for *a*-prefixing. Does this indicate that they have the same constraints as the Appalachian speakers in their grammars, rendering them polylectal? Wolfram is

doubtful that we should draw fully polylectal conclusions from this, partly because he shows that the Washington, D.C speakers do *not* have similar knowledge of the constraints on distributive *be* (found in African-American Vernacular English), but also because some of the constraints on *a*-prefixing may be simply a facet of more general principles in English (such as the impetus to avoid two neighbouring unstressed syllables, which would be produced if forms like (5) were grammatical), but this cannot account for the entirety of the extremely close match between the Washington, D.C. speakers' and the Appalachian English speakers' intuitions, which may well thus be evidence that speakers grammars can be influenced by varieties other than their own – it could be that a speaker's I-language could become polylectal through incorporating linguistic forms into their grammar which they hear, but do not use.

Polylectalism is still widely derided in most work which considers it, however. Harris (1984, 1994), Chambers & Trudgill (1980, 1998), McMahon (1992, 2000), Lodge (2009) are all critical of the idea that theoretical linguistics should entertain polylectal grammars. As well as the problems that panlectalism faces in the light of massive absolute neutralisation, it can be shown that the some of the predictions that it makes in terms of the intuitions that speakers should have about varieties of their language other than their own are not borne out – for example, non-rhotic speakers who mimic rhotic speakers often insert an [r] in places where no rhotic speaker would have it, pronouncing *last* as [laːɹst] and *father* as [fɑːɹðə], for example (Chambers & Trudgill 1980). It could be that any knowledge that speakers have of other grammars is stored elsewhere – outside of their own I-language. It is also the case that certain linguistic phenomena in English that seem to be amenable to a polylectal analysis, because it unifies things that function in the same way in different dialects, are really perceived as the same because all languages work in a particular way, rather than just all varieties of English. Thus, the reason why [ɹ, ɻ, ɾ, ʁ, ʋ] should all be analysed as realisations of the one English underlying rhotic could be not because all varieties of English work like that, but rather because that is what would be expected in any language. This would remove the support that the example might give for the polylectal approach, because the analysis could not linguistically be different.

Most work in theoretical linguistics is avowedly dialectal, and most analysts would likely claim to be idiolectal if the issue arose, doubtless agreeing with Newmeyer (2005, 230) when he writes "I had never heard a 'positive *anymore*' sentence like *We go there anymore* (in the meaning 'Nowadays we go there') until I was away at college, and the first time I heard one I couldn't parse it. How could my grammar at age sixteen conceivably have been the same as that of an English speaker to whom the construction was native?" Not all linguistic work concurs, however. Newmeyer is criticising an approach which can be seen as positively embracing polylectalism: corpus linguistics. This approach typically works with speech from more than one individual in its search for linguistic generalisations. Clark (2005, 208), who Newmeyer is attacking in the above quotation, writes that "descriptions of language use drawn from multi-individual corpora may seem irrelevant to grammar models. However, if we make the idealization that speakers share the same mental grammar, we can use the frequency of data of multi-individual corpora as a model of individual variation." This shows an approach which, while it might want to be idiolectal (describing 'individual variation'), actually embraces polylectalism in the idealisation of its data source.

It has been shown, however, that there are cases where not all speakers have the same grammar for the same feature, so Clark's approach, and polylectalism in general, might be on shaky ground. Nevins & Vaux (2003) show this for English *shm*-reduplication, which has a clear grammar for certain canonical cases that are widely known, such as *oedipus-schmoedipus*, *fancy-schmancy*, *bagel-shmagel*, removing a consonant if there is one, adding *shm-* and reduplicating the remainder of the base. English speakers have abstracted a grammar for *shm*-reduplication from this, but speakers' grammars differ when they go beyond the types of cases known from canonical forms. For example, when the base has a complex onset, both *broom-shmoom* and *broom-shmroom* were chosen by around 40% of speakers in a grammaticality test by Nevins & Vaux, and where the base begins with *shm-*, as in *shmuck*, speakers do a wide range of things: 67% claim that it cannot be reduplicated, but 27% supply a range of alternative overwriting strategies, including *shluck, shnuck, fluck, shpuck* as the reduplicated form. If this shows that speakers can have very different grammars for one particular feature, it makes it seem possible that they could have different grammars for less peripheral linguistic features, making it less reasonable to perceive them as sharing a grammar, on a polylectal approach.

[5] These issues are not restricted to English: in his introduction to a volume on the phonology of Italian, Krämer (2009) writes: "As a consequence of the regional variation, the formal analyses which will be developed in the following chapters, taken together, are not intended to represent the linguistic

competence of one individual speaker, but are intended as a 'meta-Italian' grammar, which covers a wide range of phenomena to be found in the linguistic space that can be regarded as Italian." Such an enterprise can only make sense if there is a linguistic reason to group together 'Italian' or 'English' in this way.

[6] Vaux (2008) argues that a rule based approach, with optional rules, is necessary to account for the variation found in tapping in words such as *marketability*, where either or neither /t/ can be tapped completely independently from each other (so an all-or-nothing approach, as would be necessary in other models of phonology, such as OT, fails to predict the surface variation).

[7] There is a conceptual difference between viewing this approach as involving 'partially ordered constraints' (as in Anttila & Cho, 1998, and Coetzee & Pater, to appear) or involving 'multiple grammars' (as in Anttila, 2002, who describes it first in Kiparsky 1993). The latter view might see this as avoiding intra-grammar variation – rather than using a set of constraints which are variably ranked, it uses a set of grammars involving the same constraints, but ranked differently. However, as van Oostendorp (2004, 7) explains, it is "hard to draw a distinction between these two interpretations ... If an individual has complete freedom in manipulating individual grammars, why don't we ever find a speaker who randomly chooses between pronouncing a word with a Chinese phonology and pronouncing the same word with an English phonology? We could answer this by assuming that different grammars should always be stored in a parsimonious way: redundancies between different grammars will be removed. But then we actually have something which is suspiciously similar to the intensional 'floating constraint' interpretation." Following van Oostendorp, and others, I view this approach as one which involves partially ordered constraints in a single grammar.

Additional references for the endnotes

Anttila, A. & Cho, Y. (1998) Variation and Change in Optimality Theory. *Lingua* 104. 31-56

Carr, P. & Honeybone, P. (2007) English phonology and linguistic theory: an introduction to issues, and to 'Issues in English Phonology'. *Language Sciences* 29. Special Issue on Issues in English Phonology. 117-153.

Chambers, J. & Trudgill, P. (1980) *Dialectology*. Cambridge: Cambridge University Press

Chambers, J. & Trudgill, P. (1998) *Dialectology*. Second edition. Cambridge: Cambridge University Press

Cheshire, J. (1982) *Variation in an English Dialect: a Sociolinguistic Study*. Cambridge: Cambridge University Press.

Clark, B. (2005) On Stochastic Grammar. *Language* 81. 207-217

Jackendoff, R. 2002 Foundations of Language: Brain, Meaning, Grammar, Evolution. Oxford : Oxford University Press.

Kiparsky, P. (1993) Variable Rules. Handout distributed at the Rutgers Optimality Workshop.

Krämer, M. (2009) *The Phonology of Italian*. Oxford: Oxford University Press.

Nevins, A. & Vaux, B. (2003) Metalinguistic, shmetalinguistic: the phonology of *shm*-reduplication. Proceedings of the 39th Chicago Linguistics Society. 702-721

Oostendorp, M. van (2004) Variation in phonology. Theory and Evidence. Manuscript. Available at http://www.vanoostendorp.nl/fonologie.php?lan=en

Preston, D. (1989) *Perceptual Dialectology: Nonlinguists' Views of Areal Linguistics*. Dordrecht: Foris

Wolfram, W. (1982) Language Knowledge and Other Dialects. *American Speech* 57. 3-18.