The no man’s land between syntax and variationist sociolinguistics: the case of idiolectal variability

Leonie Cornips (Meertens Instituut/Maastricht University)

Abstract: The aim of this paper is to focus on the so-called no man’s land where sociolinguistics and grammatical theory interact. It is argued that E-language as a social and I-language as a psychological construct do not exist independently, but influence each other. In other words, syntactic variation and change are driven by social factors but constrained by the nature of possible grammars. The interaction between the social meanings of linguistic forms on the one hand and grammar on the other brings about complex and multi-layered relationships between the individual and the group’s or societal grammar. This paper emphasizes how individuals are restricted by grammar but, at the same time, able to overcome these restrictions in specific situated contexts through interactions. This combined approach enables us to predict why some structures are more resistant or vulnerable to syntactic variation and change than others and the route(s) individuals may take to overcome syntactic “restrictions”. In this process of interdependent relations between the I- and E-languages, the interpretation and evaluation of linguistic forms through interaction is of crucial importance in the realization of so-called “impossible” or “unrealized” constructions.

1. Introduction

According to the editors of this volume – Aria Adli, Marco García García and Göz Kaufmann – the system–usage issue remains controversial in modern linguistics, and a single axis with the endpoints “system-based” and “usage-based” does not do justice to the complexity of linguistic reality. In this paper, I will try to go beyond the system–usage dichotomy perspective by viewing the phenomenon of syntactic variation as a crossroad where sociolinguistic and generative grammar meet. Syntactic variation could be considered a multilayered phenomenon that is the result of cognitive capacities (cf. Kayne 1994) and is strongly influenced by social or linguistic practices (cf. Labov 1972, 1994; but more specifically Eckert 2000, 2008; Silverstein 1985). The attempt to let these two approaches interact with each other (Adger 2006; Adger and Trousdale 2007; Adger and Smith 2010; Cornips and Corrigan 2005a, 2005b; Wilson and Henry 1998, among many others) should give us insight into how syntactic variation is driven by social factors but constrained by the nature of possible grammars (cf. Wilson and Henry 1998: 82).

The aim of this paper is to emphasize that there are two crucial issues regarding linguistic complexity that must be tackled within the grammar and usage debate: the issue of idiolectal or intraspeaker variation, and the complex and multilayered relationships between the individual and the society that bring about mutual effects on both the individual and the group grammar. In my opinion it is at the level of individual speakers where we can best examine the locus and limits of syntactic variation. It is here that we encounter the largest possible variation space and

1 Aria Adli suggested the first part of this title to me.
its boundaries for syntactic variation and that we find answers to the questions of whether, why and how speakers cross these boundaries.

I will present four case studies that will deepen our insight into why and how speakers in situated contexts cross syntactic boundaries, showing complex and multilayered relationships between the individual and the group grammar. These will show that speakers do not only differ from each other but that they show intraindividual variation as well. Speakers select and incessantly combine linguistic forms producing multilayered clusters of linguistic elements in social and geographical space, but this only comes about because speakers (re)interpret these forms continuously (Eckert 2008: 463). Thus, linguistic forms are always available for reinterpretation and carry various social meanings through discourse. However, the process of reinterpretation and, hence, giving social meaning to linguistic forms does not entail that speakers can select and combine linguistic forms randomly or arbitrarily. However, it does imply that ungrammatical or unacceptable constructions may become acceptable and realized in the process of interaction since speakers make meaning out of producing and interpreting linguistic forms together.

The first case study, which focuses on word order alternations in a particular three-verb cluster in Dutch dialects, shows that speakers prefer more than one order in this type of clusters and to various degrees. However, the specific combinations of the possible word orders are not distributed randomly – that is, there are combinations that are categorically absent while others are distributed but restricted in geographical space. The second case study presents variation between overt and null forms of the determiner (D). It reveals to what extent the individual speaker may accommodate and identify with his/her interlocutor resulting in intraindividual variation without any noticeable effort. The third case study deals with contemporary urban vernaculars and is particularly interesting since syntactic restrictions are loosened in these varieties. As a result, new types of constructions emerge as an integral part of grammar even if the phenomenon in question does not seem to allow for variation, as in the case of verb-second word order in Continental West Germanic varieties. Youngsters in large cities may override verb-second word order in situated contexts, but again this variation is not random since the syntactic route to overcoming it is the same for every speaker. Finally, the case study on the overuse of common gender determiners in Dutch shows that language is not a neutral medium for communication, but rather a medium through which social acts are accomplished. A speaker using common definite determiners instead of the required standard neuter one explicitly argues that he cannot use the standard form because it would make a dumb impression on his peers.

The paper is organized as follows: The second and third sections present a brief overview of the history of the generativist and sociolinguistic approaches to variation and tackle the question of why they diverged and why they seem to be converging now. These sections are important since generativist and sociolinguistic theories have important interfaces explaining why not all linguistic forms vary to the same degree (grammar component), and why speakers may cross ‘boundaries’ regarding so-called ungrammatical or unacceptable constructions, resulting in language variation and change (sociolinguistic component). The fourth section elaborates on this interplay between I- and E-languages; the fifth section presents the empirical case studies of language use illustrating the intriguing interplay between linguistic and social organization.
2. Two diverging approaches to syntactic variation

Cornips and Corrigan (2005a, 2005b) were of course not the first to point out that researchers who espouse the frameworks encapsulated by the umbrella terms “grammar” as in generative theory and “usage” as in quantitative or variationist sociolinguistics diverge quite rigidly in terms of both their methodological approaches and their theoretical persuasions. Although certain formal resonances between the paradigms have endured since the early days of their inception, the fundamental differences between them created a schism that has persisted through most of the later twentieth century (see references in Cornips and Corrigan 2005a, 2005b).

In the 1960s both sociolinguistic and formal syntax models contained formal rules that could be applied obligatorily or optionally. Formal rules in the earliest Chomskian model were transformations that connected “deep” structures with “surface” structures on the basis of rewrite rules. Optional rules, for example, derived passive, negative or question sentences from declarative sentences. Labov introduced the concept of the variable rule as an extension of this optional rule to include social and stylistic dimensions of language use along with linguistic dimensions. However, both paradigms soon followed their own avenues that is, the successive transformational models assume the existence of categorical rules only, while variationist sociolinguistics has maintained the notion of the optional rule. The two perspectives on the nature of formal rules reflect deep-seated differences between the two models. Variationists claim that the output of a linguistic rule can be probabilistic rather than discrete and that a linguistic constraint can have a quantitative rather than deterministic effect on the outcome of the process.

However, Labov based his research on the generative linguistic model by putting forward the variable rule as a means of accommodating interspeaker and intraspeaker variation. The variationist sociolinguistic practice that has evolved from studies of language variation and change since then assumes the principle of accountability as a given (cf. Sankoff 1990: 296). This principle states that variants belonging to the same syntactic (linguistic) variable must be specified by the total number of occurrences and the potential occurrences or non-occurrences in the variable environment, i.e. it ranges between 0% and 100% (cf. references in Cornips and Corrigan 2005b). This guarantees that the entire range of variable and categorical occurrences present in the data are taken into account. The notion of the syntactic variable as a structural unit and the question of which variants belong to this unit were based on the earliest generative assumption that variants have an identical underlying structure or representation which is subject to variable surface realizations (Winford 1996: 177). The alternation between active and passive constructions is an example of two different surface manifestations of the same underlying or deep structure – that is, two variants belonging to the same sociolinguistic variable (variable rule). The definition of the syntactic variable as a structural unit inevitably follows the synonymy principle. This principle is the prerequisite for variants to be assigned to the same linguistic variable; in other words, only syntactic variants that are equivocal with regard to referential meaning, i.e. variants that are “alternate ways of saying ‘the same’ thing (Labov 1972: 118)” belong to the same variable. However, the assignment of meaning or function of syntactic variants was considered problematic soon after its introduction (Lavandera 1978). Moreover, it has been suggested that some types of morphosyntactic variation rarely serve to differentiate social groups.

---

2 For a more extensive overview see Cornips (2005).
because of their dependence on pragmatic and semantic conditioning (cf. references in Cornips and Corrigan 2005b).

In the meantime, in subsequent generative models, the idea of a derivational model was abandoned in favor of a configurational model (most recently Minimalism) in which a single representation is subject to various constraints. Formal syntax postulates a blind, deterministic application of a series of procedures given a certain starting point (Chomsky 1995). As a consequence, the original notion of the linguistic variable in sociolinguistics as a structural unit has been lost or, to be more specific, syntactic variants are now considered to be different surface realizations of different underlying structures. Moreover, as theory has progressed, semantic interpretation is thought to be the “result” of the output of transformations rather than the input. The semantic component interprets the structures provided by autonomous syntax (cf. Adger and Trousdale 2007: 269–270). Given this, the semantic component can be sensitive to different structures provided by grammar, and also within formal syntax it is a question of whether two structures or two variants have the same meaning (cf. Adger and Trousdale 2007: 269–270). In fact, semantic knowledge is mediated by syntax. Since the consecutive models in generative grammar follow each other at rather high speed, different makeups of the structural unit have been detected in the last decades. Nevertheless, it is certain that as Sells et al. (1996: 173) puts it: “Variation theory needs grammatical theory because a satisfactory grammatical characterization of a variable is a pre-requisite to decisions about what to count and how to count it, and it is an essential element in the larger question about where variation is located in speakers’ grammars”.

3. Needed: reconciling approaches to account for linguistic complexity

The next two sections form the backbone of this paper. Here, it is argued that an ideal speaker-hearer provides us with a decontextualized view on syntactic variation and change, while a speaker-hearer relationship is intrinsically social by nature. We need both approaches to account for the fact that intra- and interindividual variation do not constitute a rare phenomenon but occur in normal, daily situations. A combined approach is also needed in order to account for the fact that syntactic variation and change happens all the time. A first task for a combined approach of sociolinguistics and theoretical syntax to tackle linguistic complexity would be to find out: (i) which grammatical considerations are relevant to the definition of syntactic variables; (ii) whether grammatical theory can predict the differential vulnerability of diverse linguistic forms for variation and change; and (iii) how specific situated contexts influence individual grammars (Eckert 2000; Eckert and McConnel-Ginet 1992, 1999; Meyerhoff 2002).

Of course, with respect to theoretical syntax and sociolinguistics, Wilson and Henry (1998: 2) already noted that “there have been few real attempts to marry these seemingly divergent positions” and Meechan and Foley (1994: 63) likewise suggest that the two fields “rarely, if ever, cross paths”. However, since the nineties and especially after the turn of the millennium, sincere attempts have been undertaken to integrate grammar and usage.\(^3\) Wilson and Henry (1998: 82) argue that syntactic variation is driven by social factors but constrained by the nature of possible

---

\(^3\) See brief overviews in Adger (2006); Adger and Trousdale (2007); and Cornips and Corrigan (2005a).
grammars. Cornips and Corrigan (2005a) emphasize that the generative approach has much to gain from a perspective in which the organization of the grammar is seen as somehow reflected in patterns of usage. Moreover, usage may lend strong support to a structural analysis and usage may reveal “a glimpse of grammatical structure” (Meechan and Foley 1994: 82). The core findings of sociolinguistics are that usage data are not amorphous. Rather, as Guy (2005: 563) notes, “linguistic diversity is well mannered and orderly, following observable principles of social and linguistic organization” and “[g]rammar and usage both exhibit structure and order, some of it probabilistic, some of it categorical”. The central point here is that intuitions and usage data differ, but overlap extensively as well. Usage data differ from intuitions in that they do not provide direct insight into which syntactic variants are ungrammatical. Moreover, they do not necessarily occur in the contexts that are relevant for specific theoretical concerns. On the other hand, generative practices also demonstrate that intuitions between individuals who claim to speak the same variety may differ as well and that those individual intuitions may differ with respect to context as well over time. Further, usage data coincide with intuitions to a large extent and, even more relevantly, they may provide entirely new phenomena that were not heard of before and were not predicted by theory. Consequently, theory is enhanced by analyzing and accounting for non-expected data and patterns (see for example Jensen and Christensen forthcoming).

In fact, Adger (2006) proposes to account for variation in agreement patterns not only formally, but also functionally: Variation is not only restricted to linguistic representation but also related to language in use. In his view, variation in agreement patterns is ultimately a matter of the properties of the lexicon of functional categories. The variants that make up the linguistic variable are “not simply determined by the linguistic context in which they appear, nor are they simply in free variation. Rather, they are more or less likely to be selected depending on the previous discourse, the speaker, the audience, and other psycholinguistics and sociolinguistics factors” (Adger and Trousdale 2007: 268–269). Adger and Smith (2010: 1109) argue that “the variability found in an individual speaker is two-dimensional: it may involve varying featural specification of functional categories and/or underspecification in the mapping between these categories and between morphological forms; the former modeling the kind of variation usually thought of as ‘parametric’ and the latter modeling the kind of variation usually captured by the notion of linguistic variable.” In the lexical feature model, despite its limited scope, the most recent formal insights are very well articulated. However, although syntax is viewed as completely autonomous, sociolinguistic research shows that syntax continuously varies and changes not only between generations but also at the level of the individual speaker. The ways in which individuals speak are constrained by grammar, thus syntactic variation is certainly limited and some parts of grammar are more resistant to variation than others (for instance, V2 in Germanic languages). But the other side of the story is that individuals are able to overcome so-called syntactic restrictions. In order to detect this, however, we need to study these individuals in situated contexts in interaction with others.

These contexts show how individuals divide up the world in which they live and how these oppositions obtain their shape linguistically. According to Eckert (2008, 2012), linguistic variation constitutes a robust social semiotic system that is able potentially to express the full range of social concerns in a given community; variation does not simply reflect, but constructs social meaning, hence it is a force in social change. Speakers recognize syntactic variants as stereotypes and these may be
activated (or avoided) in public performances or otherwise in highly stylized uses of local-sounding speech (Eckert 2000; Rampton 1995). Speakers do not simply reflect “grammars” or “social categories” but are agents as well. Consequently, the question of which linguistic element(s) will become socially meaningful is dependent on the individual and wider societal, political and ideological context of interaction (Cornips forthcoming). This may lead to individuals crossing or stretching the borders of the variation space of syntax in specific situated contexts and social practices.

4. Linguistic complexity: intraspeaker or idiolectal variation

Syntactic variation at the level of the individual speaker and the community is not chaotic and distributed randomly but is governed by social rules (Labov 1972, 1994 and many others). Important questions to be addressed are therefore: How and to what extent do individual and group grammars influence each other? Is it possible to predict which linguistic variants will become socially meaningful in specific situated contexts and which lexical features will be spelled out ultimately?

The relation between an individual and societal “grammar” is not one of a simple dichotomy between I- versus E-language. I-languages are the product of genetic endowment and individual experience (Chomsky 1995). People who live closely together can understand each other because they share a common genetic endowment (by virtue of being human) and a common (linguistic) experience. This experience, however, is not completely identical and, therefore, one will always find some variation between the I-languages of people who claim to speak the same dialect (see Adger and Trousdale 2007: 271). This corresponds to Guy's (2005: 562) summary of the principal findings of sociolinguistics in which he states that individuals are grammatically (more or less) similar due to social proximity. Hence, experience cannot take place in a social vacuum: It is shaped in interaction with and by others. Consequently, the opposition between I-language and E-language is not as watertight as suggested in the literature. According to Muysken (1999: 72, 2000: 41–43), the cognitive abilities which shape the I-language determine the constraints found in E-language. Moreover, the norms created within E-language make I-language coherent. Thus, it is untenable to picture “grammar” as merely a transparent representation of inner mental events since language is one of the most important mediums through which social acts, including linguistic norms, are accomplished. Language use itself is a form of social activity (Widdicombe and Wooffitt 1995: 1).

Henry (2005) has already pleaded for more attention to the phenomenon of idiolectal (intraspeaker) variation by both sociolinguists and theoretical syntacticians. Within the classical grammar perspective, variation refers to structural differences between individual grammars (interspeaker, cross-linguistic variation or variation between closely related dialects) but not within the individual grammar itself. Central questions in current syntactic research are: (i) What are the limits of syntactic variation for the individual speaker and in general? and (ii) What is the locus of syntactic variation in the grammar model? In my opinion, intraspeaker variation is the most challenging kind of variation to examine, and therefore this type of variation may be the key to gaining generative insights and answering these questions. After sixty years of theory development about the internal organization of grammar, the idealized speaker-hearer environment should be left aside and, instead, generative insights should be tested in the realm of language use where this internal organization has its most complex output. Hence, E-language as a social and I-language as a
psychological construct do not exist independently of one another, but their interaction influences the grammars of speakers and the way they speak. Hence, the multilayered relationships between language as a social product and language as “grammar” continuously influence language norms. These norms are crucial since they determine which linguistic elements are selected (or not selected) by speakers in which contexts and, consequently, relate to the central question of how people use language in their daily lives (social practices) and how grammar is organized. The norms, the selection of linguistic elements and the daily practices of people influence one another continuously (see the “total linguistic fact” by Silverstein 1985). This “holistic” view of language is the only one that can explain how individual grammars are restricted and at the same time how individuals are able to overcome these restrictions in specific situated contexts. This combined approach enables us to predict why some structures are more resistant to syntactic variation and change than others and the route(s) individuals may take to overcome these syntactic “restrictions”. In this process, the interpretation and evaluation of linguistic forms through interaction is of crucial importance in the acceptance of so-called “impossible” or “unrealized” constructions.

Very likely, the domain of micro-parametric variation – the syntactic differences between closely related individual grammars in social and geographical proximity – is the most eligible domain for addressing the questions raised above. Micro-parametric variation research clearly shows that each speaker has his/her own grammar that minimally differs from the grammar of everybody else (Cornips and Poletto 2005; Barbiers, Cornips, and Kunst 2007). In this empirical domain, there has been a clear methodological shift in generative grammar. Intuitions or native-speaker introspections in an idealized environment used to be considered the only suitable tool in macro-parametric variation research. This often meant that the resultant analyses reflected the grammaticality judgments of the theorist who may have been unaware of the considerable degree of syntactic variation which potentially exists within the same speech community (Cornips and Corrigan 2005b). However, recent micro-parametric variation research investigating dialects in many parts of Europe has drawn on acceptability judgments of non-standard speakers and, hence, it is in this domain that generative grammarians and sociolinguists are converging.

Regarding micro-parametric variation in social space, data can be collected by the so-called sociolinguistic interview – that is, systematic recordings of conversations between individuals. Due to this, analysis will consist of socially situated language samples. Of course, the setting of the sociolinguistic interview is an experimental one (Labov 1972, 1975, 1994). The data for geographical micro-parametric variation, which is the object of large-scale syntactic dialect atlases such as the Syntactic Atlas of the Dutch Dialects (acronym SAND, cf. Barbiers, Cornips, and Kunst 2007), the Northern Italian Syntactic Dialect Atlas (acronym ASIS; cf. Poletto and Benincà 2007) and the Scandinavian Dialect Atlas (acronym ScanDiaSyn; cf. Vangnes 2007) must be systematically elicited from a sample of community members (in a large geographical area) rather than derived from linguists’ own introspections. This not only enhances the empirical basis of syntactic theory, but also reduces the influence of prescriptive rules. The elicitation methodology in micro-parametric variation research relies on prior knowledge of variability within the speech community gained by observational methods, and it is on this basis that

---

hypotheses are formulated and tested. In the domain of micro-parametric variation, systematic recordings of spontaneous speech and eliciting acceptability judgments of speakers are both necessary and complementary.

5. Idiolectal variation: four case studies

In this section, I will discuss four case studies revealing intraindividual variation. The recent view on micro-parametric variation challenges the traditional idea of idiolects being sufficiently similar. Kayne (1996: XV) asks: “can anyone think of another person with whom they agree 100% of the time on syntactic judgments (even counting only sharp disagreements)? Or more precisely, are there any two people who have exactly the same syntactic judgments without exception?” (cited in Adger and Trousdale 2007: 266).

The four case studies will show that speakers differ from each other and that they show intraindividual variation even when considering word order or grammatical gender agreement. The first case study focuses on word order in the MOD-AUX-V<sub>part</sub> three-verb cluster in Dutch dialects. The second case study deals with variation between overt and null forms of the determiner (D) and reveals that the individual speaker may accommodate and identify with his/her interlocutor, resulting in intraindividual variation. The third case study shows that contact settings between youngsters in urban cities may override the so-called verb-second constraint, and the last case study shows how important intraindividual variation is for the construction of a streetwise identity.

5.1. Different (individual) word orders in the three-verb cluster MOD-AUX-V<sub>part</sub>

Let us focus first on intraspeaker variation both in elicited acceptability judgments in the SAND corpus and in recordings of spontaneous speech (cf. Cornips 1994). The SAND corpus is the result of a large-scale dialect syntax project carried out from 2000 to 2003 in the Netherlands and Belgium (cf. Cornips and Jongenburger 2001; Cornips and Poletto 2005; Barbiers, Cornips, and Kunst 2007). Different types of three-verb clusters were presented to and evaluated by 370 native speakers of local dialects throughout the Netherlands and Flanders (Belgium). Social variables of the speakers were controlled for, thus homogenizing the elicited data as much as possible with respect to age, mobility, the functional domains in which the respective dialect is spoken, and education. Only in this way it is possible to detect variation ascribed to geographical differences (cf. Barbiers, Cornips, and Kunst 2007: 60). We administered among others the three-verb cluster MOD-AUX-V<sub>part</sub> illustrated in (1):

(1)  Jan weet dat hij voor drie uur de wagen moet hebben gemaakt

‘Jan knows that he before 3 o’clock the car must have made’

The first verb in this cluster contains the modal moet ‘must’ as the hierarchically highest verb (1) (MOD), perfective hebben ‘have’ (2) as the infinitive (AUX) and the lexical verb (3), the past participle gemaakt ‘made’, as the hierarchically lowest embedded verb (V<sub>part</sub>). All six possible orders between the three verbs in (1) were offered to the subjects in an indirect relative judgment task, as illustrated in Figure 1, to collect data about variability in word orders in three-verb clusters. In this task the 370 subjects were first asked to answer with ‘yes’ or ‘no’ whether they encounter the
orders (a) through (f) in their local dialect and, subsequently, to rank the (a) through (f) from most to least acceptable on a 5-point-scale (representing *, ?*, ??, ?, ok). Thus, the subjects were instructed to consider all possible orders in Figure 1 (see also Barbiers 2005; Cornips 2009):

- a. 1-2-3 …dat…moet hebben gemaakt
- b. 1-3-2 …dat…moet gemaakt hebben
- c. 2-1-3 …dat…hebben moeten gemaakt
- d. 2-3-1 …dat…hebben gemaakt moeten
- e. 3-1-2 …dat…gemaakt moet hebben
- f. 3-2-1 …dat…gemaakt hebben moet

Table 1 presents the quantitative results of the judgment task for the whole area, i.e. the Netherlands and Dutch-speaking Belgium. It shows that the MOD-AUX-V\textsubscript{part} cluster allows four different orders (in bold) when we include only the “yes = 5” answers above a threshold of 10% (n>37) (cf. Cornips 2009). The differences in the absolute number in the columns “yes=5” through “yes=1” are due to whether the construction is acceptable in a smaller or wider geographical area. The 3-2-1 order, for example, appears to be restricted to the Frisian area, while 3-1-2 is quite common almost everywhere (in particular in the eastern part of the Netherlands) (cf. Barbiers 2005).

Table 2 shows that speakers hardly differentiate between the 2-1-3 (c) and 2-3-1 (d) order, whereas they judge the six orders representing varying degrees of acceptability (*, ?*, ??, ?, and ok) judgments. Individual speakers accept not only one but also two or even three orders. If we establish a threshold of 10% again, the MOD-AUX-V\textsubscript{part} cluster allows up to three orders per subject and the percentages of informants accepting two orders are highest, followed by informants accepting three orders:

| Table 1. Word orders in MOD-AUX-V\textsubscript{part} three-verb clusters in dialects of Dutch according to an acceptability scale of 1 through 5 (N=370, 100%) (cf. Cornips 2009) |
|----------------|----------------|----------------|----------------|----------------|----------------|
| MOD-AUX-V\textsubscript{part} moet hebben gemaakt 'must have repaired' | yes (total) | yes=5 | yes=4 | yes=3 | yes=2 | yes=1 | no, no answer |
| a. 1-2-3 | 223 = 60.3% | 107 | 47 | 37 | 20 | 12 | 147 = 39.7% |
| b. 1-3-2 | 160 = 43.2% | 78 | 32 | 33 | 9 | 8 | 210 = 56.8% |
| c. 2-1-3 | 9 = 2.4% | 4 | 2 | 0 | 1 | 2 | 361 = 97.6% |
| d. 2-3-1 | 6 = 1.6% | 2 | 1 | 0 | 1 | 2 | 364 = 98.4% |
| e. 3-1-2 | 305 = 82.4% | 210 | 49 | 30 | 9 | 7 | 65 = 17.6% |
| f. 3-2-1 | 79 = 21.3% | 44 | 14 | 15 | 5 | 1 | 291 = 78.7% |

Table 2. Idiolectal variability: the number of orders for the MOD-AUX-V\textsubscript{part} cluster accepted by the same subject regardless of the acceptability scale value (cf. Cornips 2009: 219–220)

<table>
<thead>
<tr>
<th>Type of cluster</th>
<th>1 order</th>
<th>2 orders</th>
<th>3 orders</th>
<th>4 orders</th>
<th>5 orders</th>
<th>6 orders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mod-Aux-V\textsubscript{part}</td>
<td>87 = 24.6%</td>
<td>139 = 39.4%</td>
<td>97 = 27.5%</td>
<td>26 = 7.4%</td>
<td>2 = 0.6%</td>
<td>2 = 0.6%</td>
</tr>
</tbody>
</table>
The specific combination of two, three and four orders at the level of the individual speaker is not distributed randomly. There are combinations such as the pair 1-2-3/3-2-1 that are categorically not present (n=0 in Table 3). With respect to two orders, individuals prefer the combination 1-2-3 and 3-1-2 (n= 82). With respect to three possible orders, the subjects prefer the combination 1-2-3/1-3-2/3-1-2 (n=71) significantly more than other possible three-combinations. The combination 1-2-3 / 1-3-2 / 3-1-2 / 3-2-1 is the most favorite among informants who accept four different orders (n=22):

Table 3. Combinations of accepted orders by individual speakers for the MOD-AUX-V<sub>part</sub> cluster (more than 4 subjects, cf. Cornips 2009: 220)

<table>
<thead>
<tr>
<th>two orders</th>
<th>COMBINATIONS OF ORDERS</th>
<th>MOD-AUX-V&lt;sub&gt;part&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-2-3 / 1-3-2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1-2-3 / 3-1-2</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>1-2-3 / 3-2-1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1-3-2 / 3-1-2</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>3-1-2 / 3-2-1</td>
<td>14</td>
</tr>
<tr>
<td>three orders</td>
<td>1-2-3 / 1-3-2 / 3-1-2</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>1-2-3 / 2-3-1 / 3-1-2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1-2-3 / 3-1-2 / 3-2-1</td>
<td>21</td>
</tr>
<tr>
<td>four orders</td>
<td>1-2-3 / 1-3-2 / 3-1-2 / 3-2-1</td>
<td>22</td>
</tr>
</tbody>
</table>

Importantly, the same preferences in Tables 1 and 3 were reflected in the spontaneous speech of 67 adult speakers in one location in the southeast Netherlands, namely the city of Heerlen (cf. Cornips 2009). These 67 speakers produced the order 3-1-2 in the MOD-AUX-V<sub>part</sub> cluster most frequently and they also produced the combination 1-2-3/1-3-2/3-1-2. In this respect, spontaneous speech data in one location converge with acceptability judgments in a large geographical area, i.e. the Netherlands and the Dutch-speaking part of Belgium.

This consistency in idiolectal variation calls for a combined “grammar” and “usage” approach. What is needed in order to account for the complexity of the data at the individual level is a principled answer to the question of why certain verb clusters and certain combinations of three-verb clusters are almost categorically absent. Moreover, we must analyze the size of the variation space, i.e. the most frequent and geographically constrained distributions of the various combinations of three-verb clusters. Hence, the order 3-2-1 in the MOD-AUX-V<sub>part</sub> cluster is most present in the northwest (Friesland) area, the combination of orders 1-3-2 / 3-1-2 primarily in Flanders, and the combination of orders 1-2-3 / 3-1-2 can be found everywhere in the Netherlands.<sup>5</sup> Thus on the one hand, the MOD-AUX-V<sub>part</sub> cluster shows a bewildering variation with respect to different word order alternations, which is a feature of linguistic complexity. On the other hand, this complexity shows regular patterns between closely related individual grammars distributed geographically and within one individual grammar; that is, they cluster together in social and geographical space. Clearly, the specific geographical distributions and specific word order combinations are the products of interactions between different and changing groups of speakers. To be more specific, both the acceptability judgments and spontaneous speech data (cf. Cornips 2009) provide such orderly heterogeneity that speakers can be considered members of sociolinguistic units (cf. Guy 2004). It is these clusters of

---

<sup>5</sup> See Map 1 at http://www.meertens.knaw.nl/cms/component/content/article/144225
individual speakers that may give rise to labels such as “dialect” or “language”, both for linguists as well as for laypersons.

In the strongest form of Minimalism (cf. Chomsky 1995), the Universal Grammar hypothesis states that syntactic variation does not exist. Apparent syntactic variation such as the different orders in the MOD-AUX-V\textsubscript{part} cluster should be reducible to the lexicon, that is, parameterization of morphosyntactic features, and to phonological form, thus, different ways to spell out one and the same syntactic structure phonologically (Barbiers, Cornips, and Kunst 2007). The “grammar” perspective should be able to explain why some combinations of orders are preferred above others (see table 3) by identifying a grammatical factor which is responsible for (i) the preferred combination of 1-2-3, 1-3-2 and 3-1-2 and (ii) the exclusion of 2-3-1 and 3-2-1 orders.\textsuperscript{6} The problematic issue, of course, is how to account for the fact that individual speakers can use several variants assuming that the different word orders all belong to one syntactic variable. This issue is related to questions within the successive generative models about the locus of syntactic variation, its restrictions and predictions. In the literature, two alternative approaches to this “choice” are suggested (Muysken 2005): Either the “choice” is put outside the grammatical mechanisms (Adger and Smith 2005, 2010; Adger 2006; Kroch 1989) or it is put inside the grammar by reintroducing optional rules (Henry 2002; Wilson and Henry 1998). In the same vein as Henry (1995), Barbiers (2005) argues that the categorical absence or presence of specific word orders in the verb clusters discussed above are the result of optional movement in the syntactic component. An account in terms of optional movement implies that various word orders differ only superficially from each other (cf. Harris 1996: 32; Winford 1996; Cornips 1998). In a combined approach, it should be examined how groups of speakers organize different word orders and different combinations of word orders (possibly in combination with other linguistic sets) as resources in ways that make sense for them under specific social conditions (Jørgensen 2008: 167). Accordingly, this type of variation should also be associated with particular “dialects”, social distinctions and values (Eckert 2000). Such an approach will give us a first glimpse into why inter- and intraspeaker variation is so overwhelmingly present. Such a combined approach “challenges the separation of ‘variety’ and ‘practice’ approaches” (Rampton to appear). Rampton insists that we look for the connections between dialect systems, reflexivity and interaction. The selective targeting, isolation and formal description of linguistic features remains an essential analytical task, but if we are able to construe these features as ingredients in a style or register, then we need to attend to the ways in which, with varying levels of awareness, their interactional use contributes to participants’ agentive self-positioning in the social world, aligning them with certain ideological typifications of language, speech and ways of being and not with others (cf. references cited in Rampton to appear). Let us now turn to a second case study which shows that acceptability judgments differ at the level of the individual speaker.

5.2. Acceptability judgments and idiolectal variability

\textsuperscript{6} In Cornips (2009), I argued that the 1-2-3 order as a basic structure is capable of accounting for the empirical results. The word orders in the MOD-AUX-V\textsubscript{part} cluster and their co-occurrences show that verb raising of the participle, resulting in 1-3-2 and 3-1-2 order, depends on the verbal or adjectival status of the participle (see Den Besten and Broekhuis 1989).
In this section, variation between overt and null forms of a functional head (D) will be considered. This phenomenon is of course of a very different nature than word order alternations in verbal clusters. However, both types of variation show inter- and intraspeaker variation both in written elicitation (verbal cluster described above) and in oral interviews. In the phase of oral elicitation within the Dutch syntactic atlas project (SAND), 250 locations were selected throughout the Netherlands. We had a major problem in doing the fieldwork since the large majority of the fieldworkers and Ph.D. students only speak the standard variety (cf. Cornips 2006). Therefore, we let the subject recruit an acquaintance as an “assistant interviewer” in order to be able to interview him/her in his/her own dialect. This assistant interviewer was asked to translate a standard Dutch structured elicitation task into his or her local dialect. These translations were recorded. In a second session these recordings were played to the original subject. In this second session, the entire conversation was restricted to the two dialect speakers and the fieldworker did not interfere.

One of the locations involved in the oral interviews was Nieuwenhagen in the southeast of Limburg province. In the local dialect, as elsewhere in that region, proper names are obligatorily preceded by the definite determiner *et* or *der* ‘the’. The presence of the definite determiner preceding a proper name, as in (2), is fully ungrammatical in standard Dutch:

(2) *et* Marie / *der* Jan *is* krank
    the Mary the Jan is ill

The recording of the first session between the standard Dutch-speaking fieldworker when the local assistant interviewer translates standard Dutch into his own dialect shows that the definite article in his translation is absent; the proper names *Wim* and *Els* appear without it, as illustrated in (3):

First session (dialect – standard Dutch interviewer)

(3) Ø Wim dach dat ich Ø Els han geprobeerd e
Ø Wim thought that I Ø Els have tried a
kado te geve
present to give
‘Wim thought that I tried to give a present to Els.’

In the second session, however, in which only the assistant interviewer interviews the dialect speaker in the local dialect, the latter utters the definite article as “required”:

Second session (dialect – dialect speaking interviewer)

(4) Der Wim menet dat ich et Els e boek
the Wim thought that I the Els a book
probeerd ha kado te geve.
tried has present to give
‘Wim thought that I tried to give Els a book as a present.’

This individual variation may be accounted for in different ways. From a “grammar” perspective, idiolectal variation may be explained away as a task effect (Cornips and Poletto 2005) or at least as a type of variation that is not considered to be the focus of the analysis. On the other hand, this type of variation may indeed be the result of “grammar” in use as discussed earlier – that is, the subject accommodates to and identifies with his interlocutor. In this social context, speakers can adjust to others without a noticeable effort. From this perspective, different situations of contact bring
about different types of linguistic variation. Thus, to investigate the locus of linguistic variation, to understand its social function and its spread throughout groups of speakers (communities of practices), these different contact situations require scrutiny. Hence, it is in the domain of specific interactions that syntactic restrictions or, put differently, the borders of variation space of grammar will be challenged and stretched and where so-called impossible or hitherto unrealized constructions will become possible and realized.

5.3. Peer group settings in multilingual contexts: beyond V2

In recent years, contemporary urban vernaculars have emerged among adolescents in multilingual settings in large cities throughout Europe (Auer and Dirim 2003; Blommaert 2011; Cornips and Nortier (eds.) 2008; Cornips and De Rooij 2013; Hewitt 1986; Jaspers 2005; Rampton 1995, 2005, to appear; Quist and Svendsen (eds.) 2010). These multilingual contact settings provide situated contexts in which syntactic restrictions are loosened and new types of syntactic variation emerge as an integral part of the grammar. Freywald et al. (to appear) describe how in urban vernaculars in big cities in Norway, Sweden and Germany the so-called verb-second constraint (V2) is overridden. Normally, in these languages (as also in Dutch and other Germanic languages) only one constituent may precede the finite verb in declarative clauses. Whenever a declarative clause begins with something other than the subject, subject-verb inversion is required, as in the German example in (5) (all examples from (5) through (7) are taken from Freywald et al. (to appear)):

(5)  a. Ich **war** gestern **im** Kino.
    ‘I was yesterday at the cinema.’

  b. Gestern **war** ich **im** Kino.
    ‘Yesterday I was at the cinema.’

  c. *Gestern ich **war** im Kino.
    ‘Yesterday I was at the cinema.’

However, in Norwegian, Swedish and German urban vernaculars, this restriction seems to be less robust. The typical appearance of what looks like a violation of the V2 constraint is the order “adverbial – subject – finite verb” (Adv-S-Vfin), as in example (6):

(6)  a. **I dag** hun **lagde** somalisk mat
    ‘Today she made somali food.’
    (Upus-corpus, Lukas)

  b. GESTern isch **war** KUdamm
    ‘Yesterday I was at the Ku’damm.’ [= short for Kurfürstendamm]
    (KiDKo, transcript MuH9WT)

  c. å **sen dom börjar** DRICKA **den**
    ‘And then they start drinking it.’
    (SUF-corpus, Cornelia (P47))
The overall picture points to a systematic pattern from both a “grammar” and a “usage” perspective. With respect to the former, the constituent that directly precedes the finite verb is almost without exception the subject and these subject constituents are in most cases pronominal. The adverbial in front of the subject together with the finite verb situate the event in time; it helps to structure the (narrative) discourse. The most common adverbials are the equivalents to ‘then’, ‘afterwards’, and ‘after this’ (see (6) and (7)). With respect to the latter, these “violations” of V2, or to be more precise, Adv-S-Vfin instances, occur only in peer conversations. They are remarkably rare or even entirely absent in interviews and written texts (cf. Freywald et al. to appear and references cited).

This phenomenon in vernacular urban Dutch is very rare. However, in the sporadic occasions in which it emerges it obeys the same syntactic restrictions as mentioned above: An adverbial and pronominal subject is placed before the finite verb and the adverbial in front of the pronominal subject is a temporal one:

(7) toen we hadden eerst twee autos
then we had first two cars

‘Then, we first had two cars (and later only one).’

(Utrecht/TCULT-corpus, Badir; 1 out of 20 tokens of potential “then-S-V”)

In sum, in urban vernacular speech in interactions, individuals stretch the boundaries of variation space by producing Adv-S-Vfin instances that are considered “ungrammatical” in declarative main clauses. The syntactic route is almost the same for every speaker. Thus, this is a clear example of linguistic variation that is driven by social interaction (contact settings between adolescents in large cities) but grammatically restricted, hence only certain types of adverbs in the first position of the main clause trigger V2 violations. Here we encounter a clear example of the mutual relationship between individual grammar and group grammar. Let us now consider another example of a mutual effect of this relationship, namely variation in gender distinction in the adnominal domain.

5.4. Peer group settings in multilingual contexts: underspecified gender feature

Dutch grammar classifies nouns into two grammatical gender categories: common and neuter. Grammatical gender is reflected in a number of agreeing elements accompanying or referring to the noun. Definite determiners are a clear case: Singular definite determiners vary morphologically according to the gender of the noun, as illustrated in Figure 2 below. Nouns that take the singular definite determiner de have common gender. Nouns that take the singular definite determiner het have neuter gender. There is no gender distinction on the singular indefinite determiner, which is een for both neuter and common nouns, or on the plural definite determiner, which is de for both genders.

<table>
<thead>
<tr>
<th>Gender of noun</th>
<th>Definite determiners</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Singular</td>
</tr>
<tr>
<td>common</td>
<td>de</td>
</tr>
<tr>
<td>neuter</td>
<td>het</td>
</tr>
</tbody>
</table>

Figure 2. Definite determiners in Dutch
The results of experimental tasks have revealed that the monolingual acquisition of the Dutch definite neuter determiner is a long process as children do not acquire the target system before the age of seven (cf. Blom, Polišenská, and Weerman 2008; Polišenska and Weerman 2008). Instead, monolingual children overgeneralize the definite determiner *de* and use it incorrectly with neuter nouns that require the definite determiner *het*. It takes bilingual children even longer to acquire grammatical neuter gender in Dutch. But although the overgeneralization of the definite determiner *de* constitutes a linguistic resource for every bilingual child (and even for monolingual acquirers, cf. Cornips and Hulk 2008), it only becomes socially meaningful in a process of ethnic and age identification (cf. Cornips 2008; Cornips and Hulk 2013). Nortier and Dorleijn (2008) point out that youngsters of Moroccan descent are learning both linguistic norms and norms of stylistic appropriateness. The overgeneralization of common gender is one example of this process, as illustrated by the following quotation from a conversation with a Moroccan informant S. from Rotterdam:

1 S: Dat is het slechte Nederlands
2 Int1: En heeft dat ook een naam?
3 S: Ja, niet echt, maar ‘t is in principe dan eh lidwoorden die gebruik je dan expres verkeerd
4 Int2: Ja ja die gebruik je dan expres verkeerd, net als-
5 S: Ja, dus
6 Int1: *Die meisje*
7 S: *Die huis* zeg ik dan. Terwijl ik weet ik bedoel ik weet heus wel dat het 9 *dat huis* is maar ‘t staat zo dom als ik dat op straat zeg, als ik zeg
8 Int2: Ja
9 S: Als ik zeg *dat huis*.
10 Int2: Ja ja
11 S: ‘t Is gewoon *die huis*. Maar als ik met jullie spreek dan wordt ‘t gewoon *dat huis*.
12 Int1: Yes, like
13 S: I then say that house [line 8: *die* *COMMON huis*]... While I know I mean I certainly know that it should actually be that house [line 9: *dat* *NEUTER huis*] but it would make a dumb impression if I would say
14 Int2: Yes
15 S: If I would say that house [line 11: *dat* *NEUTER huis*] in the street
16 Int2: Yes, yes
17 S: It is just that house [line 13: *die* *COMMON huis*]. But when I speak with you two (the authors both Dutch and middle-aged JN/MD) it is just that house [line 14: *dat* *NEUTER huis*].

“The speaker in the quotation explicitly says that he has to make errors, deviations from the standard norm, in order to be recognized as someone who is hanging out with friends” (cf. Nortier and Dorleijn 2008: 132). This quote suggests that the construction of identity takes place by means of a language which is governed by linguistic norms of the group the speaker wants to belong to and identifies with in contrast to others. 7 Subsequently, the interplay between the individual and the group

---

7 Of course, this example also raises interesting points concerning language awareness.
is visible in the syntactic output, i.e. gender markings on forms bearing gender, namely the determiner.

However, the overuse of common gender on determiners also shows up in other “languages” whose grammatical gender system is similar to Dutch – that is, languages that have (i) a distinction between neuter and common gender; (ii) little evidence for grammatical gender, with gender not being marked on the noun, (iii) robust frequency differences such that common nouns outnumber neuter nouns and (iv) very few morphophonological cues for the different gender forms of the determiner. Quist (2008) and Kotsinas (2001: 150) note that in urban vernacular speech in Copenhagen and Stockholm, respectively, neuter indefinite and definite determiner are being replaced by common determiners but not vice versa and that the overuse of common gender always takes place in peer group settings:

Urban vernacular speech (Copenhagen)

NEUTER varies with COMMON

*et job* ‘a job’ *en job*
*det der blad* ‘that magazine’ *den der blad*

urban vernacular speech (Stockholm)

NEUTER varies with COMMON

*ett bord* ‘a table’ *en bord*

The interplay between individual grammar and group grammar is in line with Widdicombe and Woofitt’s (1995: 36) claim that a mental state, which a grammar is, is socially shared and therefore common to different members in a given society or group. Cognition, in their view, is an individual’s mental reconstruction of shared physical and social environment. Therefore, the creation of norms and mental reconstruction are agentive processes. In other words, language is not a neutral medium of transmission of values, attitudes and opinions about a world of events “out there”, but rather a medium through which social acts are accomplished, thus language use is itself a form of social activity (Widdicombe and Woofitt 1995: 1–2).

It is in this no man’s land between syntax and variationist sociolinguistics – that is to say, in the border zone constituting the mutual relationships between the individual and the social – that we can find the answer to the questions of why and how society influences the individual (grammar) and vice versa.

### 6. Summary

The aim of this paper was to emphasize the point that there are two crucial issues with regard to linguistic complexity that must be tackled within the grammar and usage debate: the issue of idiolectal or intraspeaker variation, and the complex and multilayered relationships between the individual and society that bring about mutual effects on individual grammar and group grammar. It is at the level of individual speakers that we can best examine the locus and limits of syntactic variation. Here, we encounter the largest possible variation space and answers to the questions of
whether, why and how speakers even enlarge this space, and in doing so, overcoming syntactic constraints.

I have discussed word order alternations in three-verb clusters in Dutch dialects, the variable use of the definite determiner preceding proper nouns, the instances of Adv-S-V_{fin} in main declarative clauses in Germanic V2 languages and the use of common determiners instead of neuter ones in Dutch. Syntactic variation can be studied from two perspectives: a theoretical-analytical one in which grammar relates to genetic endowment and cognitive capabilities, and a social perspective in which language is a social praxis. Of course, these two sides of language can be studied in isolation from each other. However, an approach combining crucial theoretical insights and methodologies from each discipline brings us a step further in disentangling and understanding the phenomenon of new and complex patterns in language use data.

A combined approach of the two disciplines is the only one that enters into this no man’s land between syntax and sociolinguistics where variation is driven by social factors but constrained at the level of possible grammars. It will give us insight into why inter- and intraspeaker variation is so overwhelmingly present and why ungrammatical or impossible structures are realized. We get closer to finding answers to the following questions: (i) What are the limits and loci of syntactic variation? (ii) What is the reason for the fact that individual speakers do not use all possible linguistic resources in the construction of regional and social identities? (iii) How and why do speakers overcome ungrammatical constructions, i.e. cross the borders of variation space?

References


Convergence and Divergence in European Languages, 96–134. Cambridge: Cambridge University Press.


