



## A SOURCE OF PARAMETRIC VARIATION IN THE LEXICON

### 1. INTRODUCTION

An influential conjecture concerning parameters is that they can possibly be “restricted to formal features [of the lexicon]” (Chomsky 1995: 6; cf. Borer 1984: 2f).

In Rizzi (2009, 2011) such features are understood as instructions triggering one of the following syntactic actions:

- (1) a) External Merge
- b) Internal Merge (Move)
- c) Pronunciation/Non pronunciation (the latter arguably dependent on Internal Merge, Kayne 2005a)

Here I discuss what appears to be a particularly pervasive source of variation among languages in the domain of the lexicon (both functional and substantive) and consider whether and how it can be reduced to one of the above actions.

The variation can be preliminarily characterized as follows: **language A has two (or more) lexical items which correspond to just one lexical item in language B.**

### 2. FUNCTIONAL LEXICON

#### Example 1 (Zanuttini 1997: §3.3.1 and §3.3.2)

The Piedmontese northern Italian dialect of Turin has two sentential negative markers: *nen*, which is a neutral negative marker (it simply negates a certain proposition P), and *pa*, corresponding to standard Italian *mica*, which is a presuppositional negative marker (it negates a certain proposition P that the speaker believes to be presupposed in the context of utterance, with the effect of denying the correctness of such a presupposition).<sup>1</sup>

The northern Italian Valdotain dialect of Cogne, on the other hand, has only one sentential negative marker, *pa*, which covers both functions (it can be used either as a neutral or as a presuppositional negative marker).

\* cinque@unive.it

\*\* I wish to thank for their comments to an oral presentation of this work Richard Kayne, Iliyana Krapova, Rita Manzini, Henk van Riemsdijk, Luigi Rizzi, and Peter Svenonius, as well as two anonymous reviewers of *Linguistica* for their comments.

1 For a discussion of the pragmatic conditions required for such presuppositional negation to be felicitous see Cinque (1976) and Zanuttini (1997, Chapter 3).

### Example 2 (Cinque 1999: §4.19, §4.25 and 208 n57)

The two English adverbs *soon* and *early* seen in (2a) and (2b) are rendered in Italian by the single adverb *presto*, as seen in (3a-b):<sup>2</sup>

- (2) a) *He will soon have to get up.*  
b) *He will have to get up early.*
- (3) a) *Presto si dovrà alzare.*  
'He will soon have to get up.'  
b) *Si dovrà alzare presto.*  
'He will have to get up early.'

### 3. SUBSTANTIVE LEXICON

#### Example 1:

While Italian has separate lexical items to refer to 'arm' and 'hand', *braccio* and *mano* respectively, 'leg' and 'foot', *gamba* and *piede* respectively, Bulgarian uses one lexical item for both 'arm' and 'hand', *raka*,<sup>3</sup> and one lexical item for both 'leg' and 'foot', *krak*.

#### Example 2

To the distinct English lexical items *grandson/granddaughter* (i.e. male/female grandchild) and *nephew/niece* (i.e. male/female child of sibling), only one lexical item corresponds in Italian: *nipote*, for 'grandson'/'granddaughter'/'nephew'/'niece'.

The examples could easily be multiplied.

### 4. THE LOGIC UNDERLYING THIS PATTERN OF VARIATION

I take this pattern of variation not to be accidental, and to arise from the fact that the functional or substantive denotata of the two (or more) lexical items of language A which correspond to the unique lexical item of language B **share one component/feature** (while differing with respect to other components/features). Language B capitalizes on this shared component/feature. That is, the single lexical item of language B is uniquely specified for the common component/feature and left unspecified (in ways to which I return) for the differentiating components/features. Language A, on the other

2 I take these adverbs (more exactly, adverbial phrases) to belong to the functional lexicon as they appear to correspond in terms of position and interpretation to two independent aspectual projections (cf. Cinque 1999, Chapter 3).

3 For the lexicalization of 'hand' and 'arm' cross-linguistically the World Atlas of Language Structures (<http://wals.info/chapter/129>) reports that 228 languages have an identical word and 389 languages have two different words.

hand, capitalizes on the differentiating components/features. That is to say, its two (or more) lexical items corresponding to the single lexical item of language B are specified both for the shared component/feature and for the differentiating component(s)/feature(s).

Consider in this light the examples of the functional lexicon given in section 2 (§4.1) and those of the substantive lexicon given in section 3 (§4.2).

#### 4.1 The Functional Lexicon

Concerning the first example of section 2, we noted there, after Zanuttini (1997, §3.3.1 and §3.3.2), that the Piedmontese of Turin has two sentential negative markers, the presuppositional *pa* and the neutral *nen*. The two, in addition to the different pragmatic conditions that govern them, also differ in the position they occupy within the clause. The presuppositional negative marker *pa* necessarily precedes an adverb like *gia* ‘already’, while the neutral negative marker *nen* necessarily follows it:

- (4) a) *A l'è pa/\*nen gia parti.*  
 Cl Cl is neg already left  
 ‘He hasn’t already left.’

- b) *A l'avia gia nen/\*pa salutami cul di la.*  
 Cl Cl had already neg greeted-me that day there  
 ‘Already on that day he had not greeted me’

The overall order given in Zanuttini (1997,72) is thus **pa** > **gia** > **nen** (> *sempre*). In fact, the two can co-occur, in the expected order (*pa nen*), as shown in (5):

- (5) *Fa pa nen (\*nen pa) sulì.*  
 do neg neg that  
 ‘Don’t do that!’ (the assumption is that the addressee is about to do it)

The Valdotain dialect of Cogne, on the other hand, has only one negative marker, which can be used either as a presuppositional or as a neutral negation: *pa*. However, this is not merely a lexical quirk. When it is presuppositional *pa* precedes *dza* ‘already’; when it is neutral it follows *dza*:

- (6) a) *L'è pa dza parti?*  
 Cl is neg already left  
 ‘He hasn’t already left, has he?’  
 b) *I m'a dza pa saluià ce dzor lai.*  
 Cl me has already neg greeted that day there  
 ‘Already that day he didn’t greet me.’

The overall order given in Zanuttini (1997,82) is thus **pa** > **dza** > **pa** (> *toujou*).<sup>4</sup>

All of this seems to me to point to the presence of two specialized negative projections, which share a common core (**Negation of P**); one below the projection occupied by the adverb ‘already’, expressing simple Negation of P, and one above it, expressing Negation of P, where P is presupposed (denial of P). We can assume that these functional projections (like any other such projection) are built by selecting the relevant feature/head from the functional lexicon according to the selection properties of that feature.

If the lexical specifications of Piedmontese *nen* and *pa* are [Negation of P] and [Negation of P, P presupposed], respectively, each will be uniquely matched with the corresponding projection. If on the other hand the lexical specification of Valdostain *pa* is [Negation of P], with unspecified [P presupposed], then it will be able to match either projection.<sup>5</sup>

(7) a) Syntax:	F <sub>1</sub> >	gia	>	F <sub>2</sub>
	+Neg P			+Neg P
	+ P presupposed			
b) Lexicon:	Piedmontese			Valdotain
	<b>pa</b> : + Neg P			<b>pa</b> : +Neg P
	+ P presupposed			+/- P presupposed
	<b>nen</b> : + Neg P			

Consider now the second example of section 2, concerning the Italian adverb *pres-to*, which corresponds to both English *soon* and *early*. The relevant examples, (2) and (3), are repeated here as (8) and (9):

- (8) a) *He will soon have to get up.*  
 b) *He will have to get up early.*

4 It is not clear whether the two *pa* can co-occur (Raffaella Zanuttini, p.c.).

5 This requires not extending to such cases of underspecification the Aspects proposal that “each lexical entry automatically, by convention, contains the feature [-A] for every lexical category A, unless it is explicitly provided by the feature [+A]” (Chomsky 1965,111). The notion of ‘underspecification’ of syntactic features discussed here is different from the phonological and (one type of) morphological notion of ‘underspecification’ discussed in the literature; namely, that concerning those features (like the aspiration of onset stop consonants in English) that are predictable and thus can be expunged from the lexical representation and added through a rule (cf., among others, Archangeli 1984, Farkas 1990, and Steriade 1995). In the cases discussed here, the underspecified features are crucially not added (specified) at all, whether by rule or otherwise. It does bear some similarity, however, with the notion of underspecification employed in Distributed Morphology to account for cases of syncretism. According to Halle’s (1997) Subset Principle “the phonological exponent of a Vocabulary Item is inserted into a position if the item matches all or a subset of the features specified in that position.” (also see Embick and Noyer 2007, §2.4).

- (9) a) *Presto si dovrà alzare.*  
 ‘He will soon have to get up.’  
 b) *Si dovrà alzare presto.*  
 ‘He will have to get up early.’

When *presto* precedes the finite verb it is interpreted as ‘soon’. When it follows the verb it is interpreted as ‘early’.<sup>6</sup> The two *presto* can co-occur:

- (10) *Presto si dovrà alzare presto.* ‘He will soon have to get up early.’

In Cinque (1999) I had suggested that *presto*, qua ‘soon’, encodes “the fact that an event is going to take place *a short while after* some reference time” (p. 97) (cf. *Si rese/ renderà conto presto che lo stavano/stanno imbrogliando* ‘He soon realized/will soon realize that they were/are cheating him’). *Presto*, qua ‘early’, appears instead to be paraphrasable as ‘*a short time after* the beginning of a scale of waking-up (more generally: V-ing) times’ (cf. *Si è alzato presto* ‘He woke up early’). The shared core-component/feature of the two functional projections thus appears to be **a short time after x**. If the lexical specification of *presto* is [**a short time after x**] (x left unspecified), then *presto* will be able to match the two distinct functional projections: the one associated with [**in a short time after x, x a reference time**] and the other associated with [**in a short time after x, x the beginning of a scale of V-ing times**].<sup>7</sup>

Consider next the examples from the substantive lexicon mentioned in section 3.

## 4.2 Substantive Lexicon

The items of the substantive lexicon have components/features that, differently from those of the functional lexicon, do not match components/features of functional heads. Their components/features rather appear to match the categories with which we interpret/represent the world, broadly taken.

Consider the Italian – Bulgarian contrast shown in Example 1 of section 3. While Italian has two separate lexical items for ‘arm’ and ‘hand’ (*braccio* and *mano*, respectively), Bulgarian has a single lexical item, *raka*, which can refer to either ‘arm’ or ‘hand’. Similarly, while Italian has two separate lexical items for ‘leg’ and ‘foot’ (*gamba* and *piede*, respectively), Bulgarian has just one lexical item, *krak*, which can refer to either ‘leg’ or ‘foot’. I take this to suggest that Bulgarian expresses just the shared component/feature of ‘arm’ and ‘hand’ (namely, ‘upper limb’), and ‘leg’ and ‘foot’ (namely, ‘lower limb’), leaving unspecified what further differentiates ‘arm’ from ‘hand’ and ‘leg’ from ‘foot’. The separate lexical items of Italian for ‘arm’ and ‘hand’ and ‘leg’ and ‘foot’, on the other hand, in addition to specifying the shared component/

6 In English, *early* also has to follow the verb:

(i) He <\*early> got up <early>

7 Richard Kayne has suggested a similar analysis in class lectures, also proposing that the differentiating components/features are represented silently.

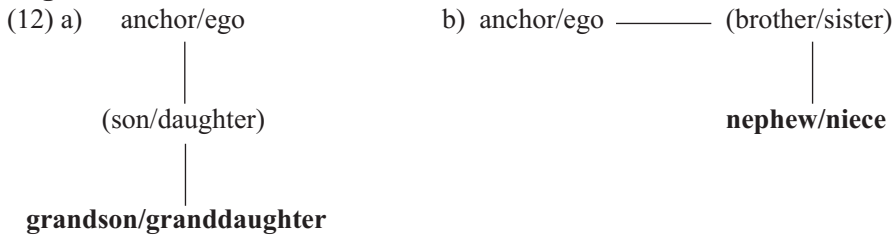
feature, also specify what differentiates ‘arm’ from ‘hand’ and ‘leg’ from ‘foot’. The lexical specifications of the different lexical items of the two languages can thus be represented in first approximation as in (11):

- (11) a) **Italian:** *braccio* ‘arm’ (+upper limb, - extremity)  
*mano* ‘hand’ (+upper limb, +extremity)  
*gamba* ‘leg’ (+lower limb, - extremity)  
*piede* ‘foot’ (+lower limb, +extremity)

- b) **Bulgarian:** *raka* ‘arm’ or ‘hand’ (+upper limb)  
*krak* ‘leg’ or ‘foot’ (+lower limb)

Consider now the second example of section 3. In Italian, a single lexical item, *nipote*, corresponds to English *grandson*, *granddaughter*, *nephew* and *niece*; abstracting away from the male/female distinction (also present in Italian in the determiners that precede the noun: *un/il* (masc.) *nipote*, *una/la* (fem.) *nipote*), *nipote* apparently corresponds in English to two distinct kinship relations, which can be represented as in (12):

**English:**



- |   |   |
|---|---|
| 1 <sup>st</sup> line: + descending, - ascending | 1 <sup>st</sup> line: - descending, - ascending (=horizontal) |
| 2 <sup>nd</sup> line: + descending, - ascending | 2 <sup>nd</sup> line: + descending, - ascending               |

These two kinship relations have, nonetheless, something in common. A degree 2 distance from the anchor/ego. The relation can be made identical if one suspends the directionality of the first line. By leaving unspecified its “descending” component/feature, one can collapse the two kinship relations into one, as in (13), which is precisely what Italian seems to do.<sup>8</sup>

<sup>8</sup> If one takes the +descending value of the first line, one gets the ‘grandson/granddaughter’ meaning; if one takes the -descending value, one takes the ‘nephew/niece’ meaning.

It is tempting to take such under-specification of components/features as a way of capturing the cross-linguistic typology of kinship terms. To mention just one example, in Western Dani (Papuan, Trans-New Guinea – Barclay 2008, 61), the lexical word *ombo* means both ‘grandparent’ and ‘grandchild’. In English *grandparent* and *grandchild* have two degrees of distance from the anchor/ego. In the former, both lines are +ascending -descending; in the latter both

**Italian: nipote**

- (13) 1<sup>st</sup> line: +/-descending, -ascending  
2<sup>nd</sup> line: +descending, -ascending

**5. UNDERSPECIFICATION VS. SILENT ELEMENTS**

In taking an ‘underspecification’ parametric approach to cross-linguistic differences in the lexicons of languages care should be taken to distinguish cases amenable to it from cases arguably involving the presence of silent elements (in one language but not in another), as in Richard Kayne’s recent work. Consider another difference between Italian and Bulgarian, which at first sight appears to be of the same ilk as the preceding ones. While Italian has one word, *molto*, for ‘(very) much’ and another word, *troppo*, for ‘too much’, Bulgarian has a single word, *mnogo*, for both. See, for example the contrast between (14), (15) and (16):

**Italian:**

- (14) a) *Non ho bevuto molto.*  
‘I didn’t drink much.’  
b) *Ha molti libri.*  
‘(S)he has got many books.’  
c) *Suo figlio è molto stupido.*  
‘His son is very (\*much) stupid.’

- (15) a) *Ho bevuto troppo.*  
‘I drank too much.’  
b) *Ha troppi libri.*  
‘(S)he has too many books.’  
c) *Suo figlio è troppo/\*molto stupido per fare una cosa del genere.*  
‘His son is too/very stupid to do such a thing.’

**Bulgarian:**

- (16) a) *Toj pie mnogo.*  
‘He drinks very much or too much.’  
b) *Toj ima mnogo knigi.*  
‘He has many or too many books.’  
c) *Sinăt mu e mnogo glupav.*  
‘His son is very or too stupid.’

---

are -ascending +descending. Western Dani *ombo* thus appears characterizable as underspecified for the +/-ascending, +/-descending components/features (provided that both lines have the same value for such components/features). This line of analysis makes us expect that no single term may cover, say, ‘grandchild’ and ‘cousin’, or ‘nephew/niece’ and ‘cousin’, or ‘grandchild’, ‘nephew/niece’ and ‘cousin’ (‘cousin’ being 3 degrees of distance: 1) + ascending, - descending; 2) - ascending, - descending; 3) - ascending, + descending).

c') *Sināt mu e mnogo glupav za da razbere tova.*  
 'His son is too stupid to understand that.' (another option is to use *tvārde* 'too (much)': *Sināt mu e tvārde glupav za da razbere tova.*)

In this case, thinking of Kayne (2005b, §3.5, 2007), there is reason to believe that the apparent ambiguity of *mnogo* 'very (much/many)/too (much/many)' is due to the presence of either one of two different silent degree words (*strašno* 'very' and *tvārde* 'too'), as these are the only degree words which are optional in the paradigms (17) and (18), and the only two which are in complementary distribution with *mnogo* in the paradigm in (19). Consider the following paradigms (Iliyana Krapova, p.c.):

- (17) a) *\*(kolko) mnogo* 'how much' (or simply *kolko*)  
 b) *\*(tolkova) mnogo* 'so much'  
 c) *\*(pò) mnogo* 'more' (or *poveče*)  
 d) *\*(naj) mnogo* 'most'  
 e) *(strašno) mnogo* 'very much/many'  
 f) *(tvārde) mnogo* 'too much/many'
- (18) a) *\*(kolko) mnogo knigi* 'how many books' (or simply *kolko knigi*)  
 b) *\*(tolkova) mnogo knigi* 'so many books'  
 c) *\*(pò) mnogo knigi* 'more books'  
 d) *\*(naj) mnogo knigi* 'most books'  
 e) *(strašno) mnogo knigi* 'very many books'  
 f) *(tvārde) mnogo knigi* 'too many books'
- (19) a) *kolko (\*mnogo) glupav* 'how stupid'  
 b) *tolkova (\*mnogo) glupav* 'so stupid'  
 c) *pò (\*mnogo) glupav* 'more stupid'  
 d) *naj (\*mnogo) glupav* 'most stupid'  
 e) *strašno (\*mnogo) glupav* 'very stupid'<sup>9</sup> *or mnogo glupav*<sup>10</sup>  
 f) *tvārde (\*mnogo) glupav* 'too stupid'<sup>11</sup> *or mnogo glupav*<sup>12</sup>

Thus the ambiguity of (16a-c) is plausibly to be attributed to the presence of a silent degree word; either *strašno* 'very' or *tvārde* 'too' (which cannot be overtly realized within an AP, if *mnogo* is). Here *mnogo* is not lexically underspecified. It acquires its apparent ambiguity as a consequence of the independent property of *strašno* 'very' and *tvārde* 'too' to be unpronounced.

9 I.e., *strašno MNOGO glupav* 'very stupid'. Capitals indicate non-pronounced elements.

10 I.e., *STRAŠNO mnogo glupav* 'very stupid'.

11 I.e., *tvārde MNOGO glupav* 'too stupid'.

12 *TVĀRDE mnogo glupav* 'too stupid'.



## 6. CONCLUSIONS

Returning now to the question posed at the beginning (whether and how the cases that we have examined so far can be reduced to one of the parametric actions seen in (1) above), it appears that while the contrast between Italian *molto/troppo* vs. Bulgarian *mnogo* is indeed amenable to the action in (1c) (pronunciation vs. non-pronunciation), the other cases examined in sections 2, 3, and 4 must be attributed to an additional parametric action: underspecification of features in the (substantive and functional) lexicon.

## References

- ARCHANGELI, Diana Bennett (1984) *Underspecification in Yawelmani phonology and morphology*. Ph.D. Dissertation. MIT.
- BARCLAY, Peter (2008) *A grammar of Western Dani*. München: Lincom Europa.
- BORER, Hagit (1984) *Parametric syntax: case studies in Semitic and Romance languages*. Dordrecht: Foris.
- CHOMSKY, Noam (1965) *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- CHOMSKY, Noam (1995) *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- CINQUE, Guglielmo (1976) “Mica.” *Annali della Facoltà di Lettere e Filosofia dell’Università di Padova* 1, 101–112. [Reprinted in: G. Cinque (1991) *Teoria linguistica e sintassi italiana*. Bologna: Il Mulino.]
- CINQUE, Guglielmo (1999) *Adverbs and Functional Heads. A Cross-linguistic Perspective*. New York: Oxford University Press.
- EMBICK, David/Rolf NOYER (2007) “Distributed Morphology and the Syntax/Morphology Interface.” In: G. Ramchand/C. Reiss (eds), *The Oxford Handbook of Linguistic Interfaces*. New York: Oxford University Press, 289–324.
- FARKAS, Donka (1990) “Two Cases of Underspecification in Morphology.” *Linguistic Inquiry* 21, 539–550.
- HALLE, Morris (1997) “Distributed Morphology: Impoverishment and Fission.” *MIT Working Papers in Linguistics* 30, 425–449.
- KAYNE, Richard (2005a) “Some Notes on Comparative Syntax, with Special Reference to English and French.” In: G. Cinque/R. Kayne (eds), *The Oxford Handbook of Comparative Syntax*, 3–69. New York: Oxford University Press.
- KAYNE, Richard (2005b) “A Note on the Syntax of Quantity in English.” In: R. Kayne (2005) *Movement and Silence*, 176–214. New York: Oxford University Press.
- KAYNE, Richard (2007) “Several, few and many.” *Lingua* 117, 832–858 (also in: R. Kayne (2010) *Comparisons and Contrasts*, 29–56. New York: Oxford University Press).
- RIZZI, Luigi (2009) “Some elements of syntactic computation.” In D. Bickerton/E. Szathmáry (eds), *Biological Foundations and Origin of Syntax*, 63–88. Cambridge, Mass.: MIT Press.

- RIZZI, Luigi (2011) "The Elements of Variation: format, locus and acquisition of parametric properties." Talk given at the *50 Years of Linguistics at MIT*, 9 Dec. 2011.
- STERIADE, Donca (1995) "Underspecification and Markedness." In: J. Goldsmith (ed.), *The Handbook of Phonological Theory*, 114–174. Oxford: Blackwell.
- ZANUTTINI, Raffaella (1997) *Negation and Clausal Structure. A Comparative Study of Romance Languages*. New York: Oxford University Press.

### Summary

#### A SOURCE OF PARAMETRIC VARIATION IN THE LEXICON

An influential conjecture concerning parameters is that they can possibly be "restricted to formal features of functional categories" (Chomsky 1995: 6). In Rizzi (2009, 2011) such features are understood as instructions triggering one of the following syntactic actions: (1) External Merge; (2) Internal Merge (Move); (3) Pronunciation/Non pronunciation (the latter arguably dependent on Internal Merge – Kayne 2005a, b). In this article I consider a particular source of parametric variation across languages in the domain of the lexicon (both functional and substantive) which appears to be due to the possibility of underspecifying certain features in some languages. The paradigmatic variation can be characterized as follows: language A has two (or more) lexical items which correspond to just one lexical item in language B.

**Keywords:** parameters, lexicon, underspecification, features

### Povzetek

#### VIR PARAMETRIČNE RAZNOLIKOSTI V LEKSIKONU

Ena izmed vplivnih domnev o naravi parametrov predvideva, da so parametri »omejeni zgolj na formalne oznake funkcijskih kategorij« (Chomsky 1995: 6). Rizzi (2009, 2011) predlaga, da te oznake služijo kot navodila jezikovni izgradnji, ki sprožijo eno od naslednjih skladijskih operacij: (1) Sestavi (External Merge); (2) Premakni (Internal Merge, Move); (3) Izgovorjava/Ne-izgovorjava (Pronunciation/Non pronunciation), pri čemer je zadnja operacija verjetno odvisna od operacije Premakni (Kayne 2005a, b). Članek obravnava specifičen vir parametrične raznolikosti v domeni leksikona (tako funkcijskega kot leksikalnega), ki je, kot kaže, posledica možnosti podspecifikacije določenih oznak v nekaterih jezikih, in pokaže, da paradigmatično variacijo lahko opišemo na sledeč način: jezik A ima dva (ali več) leksikalnih elementov, ki ustrezajo samo enemu elementu v jeziku B.

**Ključne besede:** parametri, leksikon, podspecifikacija, oznake