COORDINATION OF WORD PARTS: A SURFACE LEVEL ACCOUNT*

Ron Artstein

Rutgers University

Abstract

Coordination of parts of words is not reducible to coordination of whole words: while any group in the extension of *orthodontists and periodontists* must include at least two orthodontists and two periodontists, the denotation of *ortho and periodontists* includes groups with just one specialist of each kind. The contrast follows from a surface level interpretation of [ortho and perio]dontists. Individual word parts receive meanings through phonological decomposition: the conjuncts *ortho* and *perio* denote strings of sounds, and the remnant *dontists* is a function from sounds to word meanings that recovers the meaning of the whole word. The coordinate constituent [ortho and perio] denotes a plural object, and plural *dontists* licenses a cumulative inference, so one orthodontist and one periodontist are ortho and periodontists. The semantics also interprets ungrammatical strings like **cran and strawberries* and **peri and telescopes*, but these are ruled out on phonological grounds.

1 Introduction

In this paper I argue that the correct interpretation of coordination of parts of words, as in (1) below, requires a semantics that interprets coordination at the level of the visible string; this entails the need for separate meanings for the word parts *ortho*, *perio*, and *dontists* (an orthodontist corrects irregularities in teeth; a periodontist specializes in gums and supporting structures).

 (1) ortho and periodontists

The paper develops such a semantics, which is based on the principle of phonological decomposition (to be defined shortly); it explores the consequences this has on our understanding of the semantics of conjunction, as well as phonological constraints on coordinate structures below the word level.

The evidence that interpretation has to be at surface level comes from plural morphology. Specifically, the NP *ortho and periodontists* is not synonymous with *orthodontists and periodontists*. Suppose that Bill is an orthodontist and Martha is a periodontist; then sentence (2) below has a reading on which it is true, whereas sentence (3) does not have a true reading.

 (2) Bill and Martha are ortho and periodontists.

 (3)#Bill and Martha are orthodontists and periodontists.

The contrast between (2) and (3) is similar to the one between (4) and (5) below: only the former receives a coherent reading.

(4) Konishki and Takanohana are heavy and light sumo wrestlers.

(5)#Konishki and Takanohana are heavy sumo wrestlers and light sumo wrestlers.

The latter contrast is the result of a difference in syntactic representation: in sentence (4) the coordinate adjective *heavy and light* modifies the head noun *sumo wrestlers*; the entire NP is plural, but it can felicitously apply to a pair of one heavy sumo wrestler and one light one. The predicate in sentence (5) is formed by coordinating two plural NPs; it thus implies that each of the sumo wrestlers is both heavy and light.

Drawing on this parallel, we conclude that in (1) the conjunction *and* operates on the word parts *ortho* and *perio*. The structure relevant to interpretation must therefore be (6) below, with coordination at the surface level.

(6) [ortho and perio]dontists

The challenge, then, is to provide a semantics for word parts that will allow us to interpret structures like (6) with the correct truth conditions; this would require assigning separate meanings to the word parts *ortho*, *perio*, and *dontists*. We do want to preserve the meaning of the conjunction *and*—intuitively, it has the same meaning in (6) as it has elsewhere in the language.

I propose that the desired semantics can be formulated through the use of phonological decomposition (Artstein 2002), which derives the meanings of *ortho* and *dontist* from the meaning of *orthodontist*: in a construction like (6), the conjuncts denote strings of sound, and the rest of the word is a function from sounds to word meanings. The ordinary interpretation of conjunction then allows these denotations to combine, and results in the correct meaning for the full NP *ortho and periodontists*.

The semantics necessary for interpreting coordination of parts of words is developed in section 2. Section 3 looks at the phonological constraints on the coordination of word parts, which are important in ruling out certain configurations predicted possible by the semantics. Finally, section 4 compares the present approach to previous suggestions in the literature, which offered to explain coordination of parts of words as the result of phonological deletion.

2   A semantics for the coordination of word parts

2.1   Plurality and conjunction

Our semantics will have to capture the following difference: the NP *ortho and periodontists* can denote a pair of people, one of whom is an orthodontist and the other a periodontist, while *orthodontists and periodontists* cannot denote such a pair: it can either denote a pair of people who are each both an orthodontist and a periodontist, or a group of people of whom at least two are orthodontists and two periodontists. The source of the difference is the location of plural morphology: there is one plural morpheme on the entire NP *ortho and periodontists*, whereas in *orthodontists and periodontists* there is a plural morpheme on each conjunct.

I start by outlining some assumptions about the representation of plurality and conjunction that will be used in deriving the contrast between the above two NPs. The underlying theory of plurality and conjunction includes the following elements.

2

(7) Plurality is represented via a structured domain of individuals; plural objects are formed by a join operation $\oplus$, and are of the same type as singular individuals, namely type e. (Leonard and Goodman 1940)

(8) Plural morphology is interpreted as semantic plurality: plural expressions only include pluralities in their extension. (cf. Chierchia 1998)

(9) Coordination can receive a cumulative (plural-forming or "non-Boolean") interpretation. (Link 1983; Krifka 1990)

The above claims are defended in Artstein (2001), where the argument is made based on the behavior of coordinate adjectives in languages that mark them with plural morphology. The assumptions about plurality (7) and (8) state that plural NPs like *orthodontists* and *periodontists* denote sets of strictly plural objects.

(10) $[\![\text{orthodontists}]\!] = \lambda\alpha.\alpha \in \text{PL} \wedge \alpha = \alpha_1 \oplus \cdots \oplus \alpha_n \wedge \forall i \leq n [\![\text{orthodontist}]\!](\alpha_i)]$

$[\![\text{periodontists}]\!] = \lambda\alpha.\alpha \in \text{PL} \wedge \alpha = \alpha_1 \oplus \cdots \oplus \alpha_n \wedge \forall i \leq n [\![\text{periodontist}]\!](\alpha_i)]$

As for coordination, the assumption in (9) states that when *and* coordinates individuals of type e it can denote the join operation on the domain of individuals $\oplus$.

(11) $[\![\text{Bill and Martha}]\!] = [\![\text{Bill}]\!] \oplus [\![\text{Martha}]\!]$

Cumulative conjunction for common nouns (type et) is defined as follows: an object α is in the denotation of a coordinate common noun if it is the join of two objects α_1 and α_2, where α_1 is in the denotation of the first conjunct and α_2 is in the denotation of the second.

(12) $[\![\text{orthodontists and periodontists}]\!] =$
$\lambda\alpha.\alpha = \alpha_1 \oplus \alpha_2 \wedge [\![\text{orthodontists}]\!](\alpha_1) \wedge [\![\text{periodontists}]\!](\alpha_2)$

We thus get the desired reading for *orthodontists and periodontists*: there must be at least two orthodontists and two periodontists in any group denoted by (12); it may be the same individuals who are practitioners of both kinds, or different individuals (in the latter case the size of the group is greater than two).

2.2 Phonological decomposition

We now turn to the coordinate NP *ortho and periodontists*. In order to interpret it we must assign an interpretation to the word parts *ortho*, *perio* and *dontist*. I believe that the lexical or etymological meanings of these morphemes are largely irrelevant. Many speakers can identify the morpheme *ortho* in words like *orthodontist*, *orthopedics*, *orthography* and *orthodox*, without knowing the etymological meaning of the root and what it contributes to each of these words. What matters, then, is the ability to recognize *ortho* as part of a bigger word, whose meaning is known.

The meanings of the individual morphemes thus have to be derived from the meanings of the words they form. One possibility would be to treat the words *orthodontist* and *periodontist* as idiomatically combining expressions, a term that Nunberg *et al.* (1994) use for expressions like *pull strings*: the expression is idiomatic, yet speakers understand each of its constituents as making a distinct contribution to the idiomatic meaning. Thus, the word *strings* in the expression *pull*

strings has a meaning roughly equivalent to "connections". So while the parts have specialized meanings, the expression as a whole is compositional, and one can therefore, for example, refer to the "strings" that were being pulled. Interpreting a word like *orthodontist* as an idiomatically combining expression would involve assigning the word part *dontist* a meaning roughly equivalent to "dental specialist", and giving the word part *ortho* a modifier meaning like "teeth straightener". Of course, these meanings will be restricted to the word parts when they occur in *orthodontist*, so we do not expect to find a word like *ortholinguist meaning "a linguist who straightens teeth", just like we cannot use the word *strings* to mean "connections" outside the expression *pull strings*.

I find a number of problems with this possible way of assigning meanings to word parts. The motivation for the analysis of certain idioms as compositional expressions is the observation that the parts can also carry the special meanings on their own—they can be modified, topicalized, and referred to by a pronoun. But we do not find this behavior with word parts like *ortho*, *perio* and *dontist*, so there is less of a reason to think that these parts carry the same kind of special meanings.

Second, the analysis of Nunberg *et al.* (1994) relies on the concept of figuration, and idioms are explained as conventionalized uses of figurative language. We can see that this is true with an expression like *pull strings*: even though *pull* receives a specialized meaning in this expression, it retains the thematic structure of the standard lexical meaning of *pull*. However, it is hard to see how the concept of figuration can apply to a part of a word: it seems intuitively wrong to consider "teeth straightener" to be a figurative meaning of *ortho*, when *ortho* does not have a meaning in the first place.

My proposal is to derive the meanings for the word parts from the meanings of the complete words through phonological decomposition, in a manner similar to that proposed in Artstein (2002): the denotations of the morphemes will form a function-argument structure that, when put together, will retrieve the meanings of the original words. The singular common nouns *orthodontist* and *periodontist* denote properties of individuals (type *et*). I will assume that *ortho* and *perio* simply denote strings of sounds, which are individuals of type *e*.

(13) a. $[\![\text{ortho}]\!] \in D_e$: the string *ortho*.

  b. $[\![\text{perio}]\!] \in D_e$: the string *perio*.

That strings of sound are objects in our model, which are referred to by their own mention, is no great innovation. Roger Schwarzschild (personal communication) points out that there exist predicates that apply exclusively to such meanings, as in the sentences *ortho is disyllabic* and *perio ends in a tense vowel*. My claim is that this is the same denotation that we see in *ortho and periodontists*.

Given the denotations of *ortho* and *perio*, the semantics will have to give *dontist* a functional meaning of type *eet*, like that of a transitive verb: it will take as its first argument an object whose meaning is a string of sounds, and return the meaning of the word which is the concatenation of that string with the string *dontist*.

(14) $[\![\text{dontist}]\!] \in D_{eet}$: the function $h : D_e \to D_{et}$ such that for all $\alpha \in D_e$, $h(\alpha) = [\![\alpha\text{dontist}]\!]$ if α*dontist* is a word and $[\![\alpha\text{dontist}]\!] \in D_{et}$, undefined otherwise.

With the above definition, the composition of *dontist* with *ortho* and with *perio* yields the expected results.

(15)   a. $[\![\text{dontist}]\!]([\![\text{ortho}]\!]) = [\![\text{orthodontist}]\!]$

       b. $[\![\text{dontist}]\!]([\![\text{perio}]\!]) = [\![\text{periodontist}]\!]$

We need not worry about the fact that the function denoted by *dontist* is undefined for many objects in the model that it could take as an argument. In this respect *dontist* is like any other function: the expression *Bill-dontist* is incoherent, because it is impossible to concatenate a person with a sound in order to form a word. This is similar to what happens with an expression like *kissed democracy*, which is incoherent because democracy isn't something that can be kissed, even though it is of the right semantic type for objects of *kiss*.

2.3   Coordinating word parts

We now have the building blocks that derive the meaning of *ortho and periodontists*. Starting with the constituent *ortho and perio*, we notice that *and* operates here between two objects of type *e*, so the meaning of the coordinated constituent is the join of these two objects, that is a plural object of type *e* (just like *Bill and Martha*).

(16) $[\![\text{ortho and perio}]\!] = [\![\text{ortho}]\!] \oplus [\![\text{perio}]\!] \in D_e$

Now *dontists* has to apply to this object—it is, after all, of the right type. The meaning of plural *dontists* will be derived from the meaning of singular *dontist* by restricting its subject (the outer argument) to plurals and allowing a cumulative relation between its two arguments, as is the case with plural transitive verbs (Scha 1981). We get the following meaning for plural *dontists*.

(17) $[\![\text{dontists}]\!] = \lambda\beta\lambda\alpha.\alpha \in \text{PL} \wedge \alpha = \alpha_1 \oplus \cdots \oplus \alpha_n \wedge \beta = \beta_1 \oplus \cdots \oplus \beta_m$
$$\wedge\, \forall i \leq n\, \exists j \leq m\, [[\![\text{dontist}]\!](\alpha_i, \beta_j)]$$
$$\wedge\, \forall j \leq m\, \exists i \leq n\, [[\![\text{dontist}]\!](\alpha_i, \beta_j)]$$

Applying the meaning of *dontists* in (17) to the meaning of *ortho and perio* in (16) will give us the meaning of the NP *ortho and periodontists*.

(18) $[\![\text{dontists}]\!]([\![\text{ortho}]\!] \oplus [\![\text{perio}]\!])$
$= \lambda\alpha.\alpha \in \text{PL} \wedge \alpha = \alpha_1 \oplus \cdots \oplus \alpha_n \wedge ([\![\text{ortho}]\!] \oplus [\![\text{perio}]\!]) = \beta_1 \oplus \cdots \oplus \beta_m$
$\qquad \wedge\, \forall i \leq n\, \exists j \leq m\, [[\![\text{dontist}]\!](\alpha_i, \beta_j)] \wedge \forall j \leq m\, \exists i \leq n\, [[\![\text{dontist}]\!](\alpha_i, \beta_j)]$
$= \lambda\alpha.\alpha \in \text{PL} \wedge \alpha = \alpha_1 \oplus \cdots \oplus \alpha_n$
$\qquad \wedge\, \forall i \leq n\, [[\![\text{dontist}]\!](\alpha_i, [\![\text{ortho}]\!]) \vee [\![\text{dontist}]\!](\alpha_i, [\![\text{perio}]\!])]$
$\qquad \wedge\, \exists i \leq n\, [[\![\text{dontist}]\!](\alpha_i, [\![\text{ortho}]\!])] \wedge \exists i \leq n\, [[\![\text{dontist}]\!](\alpha_i, [\![\text{perio}]\!])]$
$= \lambda\alpha.\alpha \in \text{PL} \wedge \alpha = \alpha_1 \oplus \cdots \oplus \alpha_n$
$\qquad \wedge\, \forall i \leq n\, [[\![\text{orthodontist}]\!](\alpha_i) \vee [\![\text{periodontist}]\!](\alpha_i)]$
$\qquad \wedge\, \exists i \leq n\, [[\![\text{orthodontist}]\!](\alpha_i)] \wedge \exists i \leq n\, [[\![\text{periodontist}]\!](\alpha_i)]$

We find that *ortho and periodontists* denotes the set of all plural objects that are composed of singular individuals where each such individual is either an orthodontist or a periodontist, and at least one such individual is an orthodontist, and one is a periodontist. In particular, one such plural object is the join of the orthodontist Bill and the periodontist Martha. So our semantics succeeds in interpreting the NP *ortho and periodontists* at surface level.

5

2.4 Coordinate parts with number marking

Let's look again at the underlying cause for the contrast between the NPs *ortho and periodontists* and *orthodontists and periodontists*: the former has one plural morpheme on the entire NP, whereas the latter has a plural morpheme on each conjunct. This leads to the following expectation: if each coordinate part were to bear plural morphology, then coordination of parts should have the same meaning as coordination of whole words.

We find such a case in Hebrew. In compounds, number is usually marked on the head; number marking on the non-head—singular *sefer* and plural *šinayim* below—is irrelevant in determining the number of the compound, (cf. Borer 1988).

(19) SINGULAR   PLURAL
     bet sefer   bat-**ei** sefer   'school' (lit. 'house-book')
     rofe šinayim rof-**ei** šinayim   'dentist' (lit. 'doctor-teeth')

In a small class of compounds, however, number is marked on both parts.

(20) SINGULAR   PLURAL
     sgan aluf   sgan-**ei** aluf-**im**   'lieutenant colonel'
     tat aluf   tat-**ei** aluf-**im**   'brigadier general'

Hebrew compounds are considered to be morphological words because of their syntactic behavior, namely that they do not allow extraction (Borer 1988). As far as the semantics goes, the compounds *sgan aluf* and *tat aluf* should be considered to have atomic meanings because of their opacity. The morpheme *sgan* is also an independent word meaning 'deputy', and *tat* is a prefix with a meaning similar to 'sub-'; in combination with *aluf*, both of the compounds formed denote ranks that are below *aluf* 'major general'. However, exactly what ranks these are and how they are ordered with respect to one another is completely conventional, and does not follow from the meanings of the individual morphemes. (There is another sense in which the terms *sgan aluf* and *tat aluf* may be thought to be compositional: parallel ranks in other security forces have s similar structure. Thus, *sgan aluf*, *sgan nicav*, *sgan gondar* and *sgan tafsar* are parallel ranks, respectively, in the Israeli Defense Force, National Police, Prison Service and Firefighting Service; the same goes for *tat aluf*, *tat nicav*, *tat gondar* and *tat tafsar*. While these terms are related, I don't think this reduces their opacity.)

When number is marked on both parts of the compound, we find that coordination of parts requires that each conjunct should correspond to a plural referent, just like coordination of full compounds. Both of the sentences below require there to be at least two lieutenant colonels and two brigadier generals at the party.

(21) etmol   hayu ba-mesiba sgan-ei aluf-im   ve-tat-ei   aluf-im
     yesterday were at.the-party deputy-pl general-pl and-sub-pl general-pl
     "At the party yesterday there were lt. colonels and brigadier generals"

(22) etmol   hayu ba-mesiba sgan-ei   ve-tat-ei   aluf-im
     yesterday were at.the-party deputy-pl and-sub-pl general-pl
     "At the party yesterday there were lt. colonels and brigadier generals"

Our semantics should now explain why the above two sentences receive the same meaning. Coordination of full NPs (21) works the same way as with *orthodontists and periodontists*; coordination of parts (22) will have to take the plural marking on the conjuncts into account.

We start, as before, by noting that the coordinate parts denote strings of sounds; these sounds include the plural morphemes, which are within the coordinate parts.

(23)   a. $[\![\text{sgan-ei}]\!] \in D_e$: the string *sgan-ei*.

        b. $[\![\text{tat-ei}]\!] \in D_e$: the string *tat-ei*.

The element outside the coordinate structure receives a functional meaning of type *eet*, from sounds to common noun meanings.

(24) $[\![\text{aluf-im}]\!] \in D_{eet}$: the function $h : D_e \to D_{et}$ such that for all $\alpha \in D_e$, $h(\alpha) = [\![\alpha\text{aluf-im}]\!]$ if α*aluf-im* is a word and $[\![\alpha\text{aluf-im}]\!] \in D_{et}$, undefined otherwise.

This will combine with the previous meanings to yield the desired results.

(25)   a. $[\![\text{aluf-im}]\!]([\![\text{sgan-ei}]\!]) = [\![\text{sgan-ei aluf-im}]\!]$

        b. $[\![\text{aluf-im}]\!]([\![\text{tat-ei}]\!]) = [\![\text{tat-ei aluf-im}]\!]$

Notice how we had to use the plural *aluf-im* rather than singular *aluf* in our phonological decomposition. This is because the coordinate parts themselves contain plural morphemes, so their phonological concatenation with singular *aluf* would result in a non-word: if we had tried to combine the meanings in (23) with the *eet*-type meaning of *aluf*, we would not get a meaning that we could later build on.

(26)   a. $[\![\text{aluf}]\!]([\![\text{sgan-ei}]\!])$: Undefined (no word *sgan-ei aluf)

        b. $[\![\text{aluf}]\!]([\![\text{tat-ei}]\!])$: Undefined (no word *tat-ei aluf)

The significance of this is apparent: whereas previously we have introduced a cumulative inference through the derivation of plural *dontists* from singular *dontist*, such a move with *aluf-im* would be useless. Rather, we will have to define the cumulative inference of *aluf-im* with reference to its basic *eet*-type meaning.

Since the word part *aluf-im* is plural, it should allow a cumulative relation between its two arguments (just like *dontists*): *aluf-im* has to be closed under cumulative inference.

(27) $[\![\text{aluf-im}]\!] = \lambda\beta\lambda\alpha.\alpha \in \text{PL} \wedge \alpha = \alpha_1 \oplus \cdots \oplus \alpha_n \wedge \beta = \beta_1 \oplus \cdots \oplus \beta_m$
$$\wedge \, \forall i \leq n \, \exists j \leq m \, [[\![\text{aluf-im}]\!](\alpha_i, \beta_j)]$$
$$\wedge \, \forall j \leq m \, \exists i \leq n \, [[\![\text{aluf-im}]\!](\alpha_i, \beta_j)]$$

The coordinate constituent *sgan-ei ve-tat-ei* denotes a plural object of type *e*, just like *ortho and perio*.

(28) $[\![\text{sgan-ei ve-tat-ei}]\!] = [\![\text{sgan-ei}]\!] \oplus [\![\text{tat-ei}]\!] \in D_e$

Applying the meaning of *aluf-im* in (27) to the meaning of *sgan-ei ve-tat-ei* in (28) gives us the meaning of the NP *sgan-ei ve-tat-ei aluf-im*.

(29) $[\![\text{aluf-im}]\!]([\![\text{sgan-ei}]\!] \oplus [\![\text{tat-ei}]\!])$
$$= \lambda\alpha.\alpha \in \text{PL} \wedge \alpha = \alpha_1 \oplus \cdots \oplus \alpha_n \wedge ([\![\text{sgan-ei}]\!] \oplus [\![\text{tat-ei}]\!]) = \beta_1 \oplus \cdots \oplus \beta_m$$
$$\wedge \, \forall i \leq n \, \exists j \leq m \, [[\![\text{aluf-im}]\!](\alpha_i, \beta_j)]$$
$$\wedge \, \forall j \leq m \, \exists i \leq n \, [[\![\text{aluf-im}]\!](\alpha_i, \beta_j)]$$
$$= \lambda\alpha.\alpha \in \text{PL} \wedge \alpha = \alpha_1 \oplus \alpha_2$$
$$\wedge \, [\![\text{aluf-im}]\!](\alpha_1, [\![\text{sgan-ei}]\!]) \wedge [\![\text{aluf-im}]\!](\alpha_2, [\![\text{tat-ei}]\!])$$
$$= \lambda\alpha.\alpha \in \text{PL} \wedge \alpha = \alpha_1 \oplus \alpha_2 \wedge [\![\text{sgan-ei aluf-im}]\!](\alpha_1) \wedge [\![\text{tat-ei aluf-im}]\!](\alpha_2)$$

We find that *sgan-ei ve-tat-ei aluf-im* denotes the set of all plural objects that are the join of two objects, one of which is in the denotation of *sgan-ei aluf-im* and the other in the denotation of *tat-ei aluf-im*. Each such object will include at least two lieutenant colonels and two brigadier generals, as desired.

I now address a potential objection to my analysis above: one might claim that the NP *sgan-ei ve-tat-ei aluf-im*, with coordination of parts, receives the same interpretation as a coordination of full NPs because the conjoined elements are the heads of the compound, unlike the English example *ortho and periodontists*. I am actually not sure about the headedness of the above compounds, despite the formal similarity to construct states, and I think the fact that number is doubly marked by morphology may be an indication that the headedness relation is not that straightforward. More importantly, coordination of singular terms shows a contrast between compounds of the type we are looking at and truly headed compounds. If it were a matter of conjoining heads, one might expect that conjunction of singular heads might be plural; this is what we see, for instance, in the following sentence (as mentioned above, the plurality of *šinayim* "teeth" does not affect the number marking of the compound).

(30) etmol   hay-u ba-mesiba rofe   ve-rof-at   šinayim
     yesterday were-**pl** at.the-party doctor.m.**sg** and-doctor-f.**sg** teeth
     "At the party yesterday there were a male dentist and a female dentist"

But when we coordinate parts of a singular compound of the type discussed throughout this section, the result is still singular, whereas coordination of singular compounds is plural. We thus see a contrast between (31) and (32), where the latter is ungrammatical because the singular subject does not agree with the plural verb; coordination of parts of a singular compound is fine when it can felicitously refer to a single individual, as in (33).

(31) etmol   hay-u ba-mesiba sgan   aluf   ve-tat   aluf
     yesterday were-**pl** at.the-party deputy.**sg** general.**sg** and-sub.**sg** general.**sg**
     "At the party yesterday there were a lt. colonel and a brigadier general"

(32)*etmol   hay-u ba-mesiba sgan   ve-tat   aluf
     yesterday were-**pl** at.the-party deputy.**sg** and-sub.**sg** general.**sg**

(33) dani   hikir et matan vilnai betor sgan   ve-tat   aluf
     Danny knew acc Matan Vilnai as   deputy.**sg** and-sub.**sg** general.**sg**
     "Danny knew Matan Vilnai as a lt. colonel and a brigadier general"

This is predicted by our semantics, provided that we restrict the *eet*-type denotation of *aluf* to singularities, the same way we restricted that of *aluf-im* to pluralities.

3   Phonological constraints

Phonological decomposition allows us to interpret coordination of parts of words at surface level, yielding the correct interpretation for plural morphology on the conjuncts (or lack thereof). The semantics, however, does not make a distinction between grammatical strings like *ortho and periodontists* and ungrammatical ones like **cran and strawberries* or **peri and telescopes*. And indeed it should not: it is phonological rather than semantic factors that are responsible for the above contrast. This section explores the phonological constraints on the coordination of parts of words.

3.1 Phonology and etymology

Before I go into the details of the phonology, I want to address a concern that the contrast just mentioned may be the result of the words *orthodontist* and *periodontist* being in some sense more "compositional" than the words *cranberry* and *strawberry*. This may be true from a historical or etymological point of view; synchronically, however, there is no difference in compositionality. Speakers of English can readily identify the morphemes *-berry* and *-dontist*, since the words that contain them share some common aspects of meaning. But the morphemes *cran*, *straw*, *ortho*, and *perio* seem just about equally opaque. There is nothing more to the meaning of *cran* and *straw* other than the kind of berry they signify. The term *orthodontist* is in common use, and it seems reasonable to assume that speakers learn it as a unit; the relation to words like *orthodox* and *orthography* is fairly obscure. The word *periodontist* is much less familiar, and it is not at all easy to arrive at its meaning, despite familiarity with the morpheme *peri* in words like *perimeter* and *periphery*. It appears that synchronically, the words *orthodontist* and *periodontist* are just as opaque as *strawberry* and *cranberry*.

So where do speakers get the intuition that *orthodontist* and *periodontist* are more compositional than *strawberry* and *cranberry*? I believe this is a phonological intuition, rather than a semantic one. Some speakers are more content with *boysen and huckleberries* than with **cran and strawberries*, even though it is hard to see in what sense it is more compositional. Furthermore, speakers are happy to accept *ortho and periodontists* even when they do not know exactly what a periodontist is, simply assuming it is some kind of dental specialist. This would be hard to explain if acceptability were tied to the lexical meaning of the word part *perio*.

Another possibility is that compositionality is not an intuition about meaning, but rather an intuition about combinatorial properties: speakers are aware that the morphemes *ortho* and *peri* form a variety of compounds, whereas *cran* and *straw* are less productive in this sense (note, however, that in contemporary English, *cran* is used productively and rather transparently in compounds or blends like *cranapple* and *cran-grape*). The fact that *ortho* and *peri* belong to the scientific, learned vocabulary may also give speakers the impression that they are, or should be, more transparent, even if the speaker doesn't know why. But the combinatorial properties still do not explain the coordination facts. If it were the combinatorial properties alone that allowed coordination of parts, we should expect structures like **peri and telescopes* to be well formed, since *periscope* and *telescope* are also compositional, in the etymological sense, as much as *periodontist* and *orthodontist*. Our conclusion is that the intuition that coordination of parts is allowed by compositionality of lexical meaning or even etymological compositionality is a false intuition; the difference between the well formed examples and the ungrammatical ones is due to the phonology.

3.2 Description of the prosodic facts

Structures where word parts are coordinated are subject to certain prosodic restrictions. Our first observation is that the coordinate parts must be separated from the rest of the word by a certain prosodic juncture. Okada (1999) gives minimal contrasts such as the following (p. 350).

(34) a. *physio and psychologies
      b. physio and psychological

Semantic factors do not explain the contrast between the above examples. Nor can the contrast be attributed to morphological structure, as *physio* and *psycho* are identifiable morphemes in both

examples. Rather, the contrast is a matter of prosodic (metrical) structure: a single foot may not span segmental material from both the coordinate and non-coordinate parts (in the following example the offending foot is underlined).

(35) a.*physio and psy(ˈcholo)(ˌgies)

  b. (ˌphysio) and (ˌpsycho)(ˈlogi)cal

We can formulate a first approximation of the prosodic constraint on coordination of parts of words—material from the coordinate part and the part outside the coordinate structure may not be prosodified into one foot. An alternative way to state this is that there must be a foot boundary between the coordinate part and the rest of the word.

The above restriction still does not predict the ungrammaticality of *cran and strawberries, since cran, straw and berries can all be parsed into separate feet. I think that this is desirable. While it is true that *cran and strawberries is unacceptable to many (perhaps most) speakers, it is still markedly better than *physio and psychologies. The minimal requirement is that the coordinate and non-coordinate parts should be separated by a foot boundary; structures that satisfy this requirement have varying degrees of acceptability, and coordination of parts of words in English is dispreferred when primary stress is on the coordinate parts. This is similar to the behavior of expletive infixation (McCarthy 1982): orthobloodydontist and psychobloodylogical are better than ?cranbloodyberry, but the latter is definitely preferable to the impossible *psychobloodylogy.

Another prosodic generalization is that the free conjunct in the coordinate structure invariably forms a prosodic word. I know of two tests that can detect a prosodic word boundary in English: stem-final tensing and [ɹ] intrusion. Stem-final tensing prohibits the reduction of non-low vowels to schwa at the end of a prosodic word (Booij and Rubach 1987); intrusive [ɹ] appears in certain dialects of English between a non-diphthongal vowel and a following vowel, when the two vowels are separated by a prosodic word boundary (McCarthy 1993). The examples below—(36) for all English speakers and (37) for those speakers with [ɹ] intrusion—show that the first conjunct in the coordinate structure is parsed as a prosodic word.

(36) orth[oʷ] and periodontists (cf. orth[ə]dontist)

(37) mega[ɹ] and gigabytes

The fact that *periodontist* may be pronounced with a schwa in (36) shows that the conjunct part in that word does not have to form a prosodic word. Coordination of parts of words thus requires that the coordinate parts be constituents that can be promoted to the status of a prosodic word, and that the free-standing conjunct actually be promoted to this status (cf. Smith 2000).

Dutch and German show similar prosodic restrictions to English—there must be a foot boundary between the coordinate and non-coordinate parts of a word, and the free-standing conjunct must form a prosodic word—but they do not show the English preference against stress on the coordinate parts. Booij (1985) formulates a rule that states that coordination in Dutch and German is allowed when the element outside the coordinate structure forms a prosodic word (this is formulated in connection with a rule of phonological deletion, see section 4, but it can be equally well stated as a condition on surface representations). Prosodic word status is determined by the word formation processes of compounding and affixation. However, the examples in Booij (1985) which do not allow coordination either have the coordinate and non-coordinate parts parsed into a single foot, or have a non-coordinate part that is smaller than a full foot (I cannot reproduce the examples here due to space considerations; the reader is referred to the original article). On the

other hand, Booij gives some examples of coordinate structures like *mono- en dialogen* 'monologues and dialogues', where morphological structure does not lead us to expect that the element outside the coordinate structure is a prosodic word. If the requirement on coordination in Dutch is the same as I have suggested for English, namely that there just be a foot boundary between the coordinate and non-coordinate parts, then all the data are accounted for.

The situation in German may be more complicated—I know of at least some speakers who refuse to accept **Mono und Dialoge*, even though there is a foot boundary between the coordinate and non-coordinate parts; a foot boundary may not be enough for these speakers. Höhle (1982) gives an analysis of coordination in German in terms of morphological boundaries, but notes (fn. 8, p. 91) that often these boundaries do not correlate with other phonological or morphological tests, and may in some cases lead to speaker variation and underdetermination. A prosodic explanation to speaker variation may be that in German the minimal boundary for coordination of word parts is that of a prosodic word rather than a foot, and variation stems from different speakers having different prosodic representations for the same words. Variation may also be the result of different speakers requiring different boundaries—some speakers only allow coordination of parts of words with a prosodic word boundary, while others are content with a foot boundary. I leave the matter unresolved, pending further data.

4 Surface coordination vs. deletion

Our theory now has two parts: a semantic theory for the interpretation of coordination on parts of words, and phonological restrictions on the structures that can be coordinated. These phonological observations draw on previous analyses of similar constructions in Dutch and German, which preferred to view these constructions as resulting from a process of phonological deletion, rather than coordination at surface level (Höhle 1982; Booij 1985; Wiese 1992, 1996; Kleinhenz 1997; Smith 2000). Our motivation for an analysis of surface coordination was the difference in meaning between NPs with coordinate word parts (*ortho and periodontists*) and coordinations of full NPs (*orthodontists and periodontists*). In this section I look at additional arguments that compare surface coordination with phonological deletion.

The rule of deletion, as it is put forward in Booij (1985) and Kleinhenz (1997), includes a syntactic component as well as a phonological component. Deletion itself is phonological—the deleted element is a prosodic word. But prosodic restrictions alone do not predict the following contrasts, since in each pair the two structures are identical in terms of prosodic structure.

(38)   a. de land- en de tuinbouw   (Dutch, Booij 1985)
           'the agriculture and the horticulture'

      b.*de land- met de tuinbouw
           'the agriculture with the horticulture'

(39)   a. eine elf-   und eine zwölfjährige   (German, Kleinhenz 1997)
           an   eleven and a   twelve-year-old

      b.*eine elf-   bewundert eine zwölfjährige
           an   eleven admires   a   twelve-year-old

These examples lead Booij and Kleinhenz to conclude that the context for deletion must be syntactic: the deleted element has to be adjacent to a conjunction.

A question which is not addressed is why this particular deletion rule should exist and not, say, a rule that allowed phonological deletion, under identity, subject to adjacency to a preposition.

Indeed, if the deletion rule has nothing to do with the meaning of conjunction, then we should expect that a language should be possible where constructions like (38a) are bad, but constructions like (38b) are grammatical!

Surface coordination, as argued for in this paper, gives a straightforward answer to the above question, which follows from the meanings of parts of words. Phonological decomposition states that coordinate parts of words denote strings of sounds, which are individual objects (type *e*). These denotations can be conjoined: the conjunction of two strings of sound is a plural object which is the join of the individual strings. However, there is no way to modify a string of sound using a preposition like *with*, or to have a string of sound as the object of *admire*. It follows that parts of word are licensed in coordinate structures like (38a) and (39a), but not as objects of prepositions (38b) or verbs (39b).

The above explanation is semantic, just like the explanation for the following contrast, which does not involve parts of words.

(40) big and small monkeys

(41)*big with small monkeys (cf. big monkeys with small monkeys)

(42)*big admire small monkeys (cf. big monkeys admire small monkeys)

The adjectives *big* and *small* can be coordinated to form a constituent with a coherent meaning, which can then combine with *monkeys*. But *big* and *small* cannot combine with *with* to form a constituent **big with small*, nor can they combine with *admire* to form a constituent **big admire small*. Indeed, it would seem odd to have an analysis that derived (40) from *big monkeys and small monkeys* through deletion of *monkeys*, and then stipulated somehow that this is not possible in (41) and (42). Phonological decomposition allows the same explanation for constructions above and below the word level.

Wiese (1992, 1996) also notes an empirical problem with a deletion rule phrased in syntactic terms: it does not cover the full range of data. The following examples from German show how parts of words appear as independent syntactic elements in structures that do not involve coordination (Wiese 1996, p. 72).

(43) a. Sachsen entwickelte sich vom Herzog- zum Kurfürstentum.
        'Saxony developed from a dukedom to an electorate.'

      b. Formen wir den Aktiv- in einen Passivsatz um.
        'We form the passive sentence from the active sentence.'

      c. … übernahm zum Fraktions- auch noch den Landesvorsitz
        '… taking over both the faction chair and the state chair.'

      d. Weil Leitungs- von Mineralwasser unterschieden werden muß, …
        'since tap water must be distinguished from mineral water, …'

Wiese proposes that these structures are the product of a purely phonological deletion rule: not only is the deleted element a prosodic word (ω), but the context is also phonological. A prosodic word may be deleted at the edge of a phonological phrase (ϕ), if an identical prosodic word is contained in an adjacent phonological phrase within the same intonational phrase (I).

(44) a. $[_I [_\phi \cdots \underline{\omega_i} {}_\phi][_\phi \cdots \omega_i \cdots {}_\phi]_I]$
      b. $[_I [_\phi \cdots \omega_i \cdots {}_\phi][_\phi \underline{\omega_i} \cdots {}_\phi]_I]$

Unfortunately, this proposal fails to distinguish between the pairs of examples in (38) and (39) above. Whereas a deletion rule with a syntactically specified environment prohibits the grammatical structures in (43), a deletion rule with a prosodically specified environment allows the ungrammatical structures (38b) and (39b). Our surface coordination account captures meaning contrasts that a deletion account misses (our original motivation), and it explains why parts of words occur in coordinate structures but not as objects of *with* or transitive verbs. We still need an explanation how a surface interpretation is possible for the non-coordinate structures in (43). I do not have an answer at the moment, but we should note that these structures have parallels at the phrasal level: parts of words are allowed where parts of NPs are allowed.

(45)   a. Parliament developed from a unicameral to a bicameral institution.

b. We transform the active into the passive sentence.

c. … taking over both the faction and the state chair.

It appears, then, that however the sentences above are interpreted, this should generalize to the constructions with word parts.

Why, then, was deletion proposed in the first place? It turns out that most of these arguments rest on assumptions about the rigidity of syntax and semantics, and are no longer valid when we consider a syntax and semantics capable of manipulating and interpreting parts of words.

One argument for deletion is syntactic: sentences (46) and (47) contain apparent instances of coordination of non-constituents: *een derde* and *een zesde* in (46), and *ein Dreigang* and *ein Sechsgang* in (47).

(46) het verschil   tussen   een derde- en een zesdeklasser   (Dutch, Booij 1985)
     the difference between a   third   and a   sixth-former

(47) ein Dreigang-   und ein Sechsgangfahrrad   (German, Kleinhenz 1997)
     a   three-speed and a   six-speed-bicycle

Indeed, a syntax would need to allow some non-traditional constituents in order to generate the above examples with surface coordination. Such syntactic theories have been developed for non-traditional constituent coordination at the phrase level ("Right Node Raising": Ades and Steedman 1982; Steedman 1985, 1987), and a similar grammar is used by Moortgat (1987) for the internal structure of words.

Another argument for deletion has to do with the semantics of number. Booij (1985) notes that the preposition *tussen* 'between' in (46) requires a plural complement, whereas on a surface interpretation its complement would be interpreted as singular because of the singular head *klasser*. The same problem, however, appears with coordination of determiner-adjective sequences that are whole words.

(48) I must choose between a yellow and a red cabinet.

I do not have an explanation for (48), and perhaps a deletion analysis can be motivated for it as well as for (46). However, this is not evidence for deletion in all cases of coordination of parts of words—no more than (48) is a motivation for deletion in the NP *yellow and red cabinets*.

The next argument says that coordination of parts of words cannot be generated by the mechanism of syntactic composition because the conjuncts do not have to be free morphemes (49)–(50) or because they can be lexical items of different categories (51)–(52).

(49) schei- en natuurkunde 'chemistry and physics' (Dutch, Booij 1985)
(literally: 'analysis and nature knowledge').

(50) Him- und Brombeeren 'raspberries and blackberries' (German, Kleinhenz 1997)

(51) leer- en handboeken 'textbooks and handbooks' (Dutch, Booij 1985)
(literally: 'learn and hand books').

(52) Verband Geburts- und andere-r   Behinderte-r (German, Kleinhenz 1997)
society birth   and other-GEN handicapped-GEN

Moortgat (1987) offers a syntax for these structures: conjunct stems of various categories are converted into modifier bound morphemes through a category-changing rule. A second argument for deletion is that the lexical meanings of the above morphemes are opaque. With phonological decomposition as developed in this paper this is no longer a problem for the semantics.

The final argument for deletion comes from Dutch "linking phonemes". When *wesp* 'wasp' forms a compound with *steek* 'sting', an additional schwa [ə] appears between the two morphemes: *wespesteek*; similarly, *zonsverduistering* 'sun-eclipse' contains a linking [s]. The linking phonemes are retained when such morphemes are coordinated.

(53) wespe- en bijesteken 'wasp and bee stings' (Booij 1985)

(54) zons- en maansverduisteringen 'solar and lunar eclipses' (Booij 1985)

Booij (1985) argues that deletion is the only possible source for the linking phoneme in the first conjunct, since coordinated words are inaccessible to morphological rules: the ordinal derived from the cardinal number *drie-en-zestig* 'sixty three' is *drie-en-zestigste* 'sixty third', where the first conjunct retains its cardinal form, rather than **derde-en-zestigste* (cf. *derde* 'third'). However, there is no reason to believe that coordination should be similar to this sort of morphological derivation. Moortgat (1987, p. 47) notes that the linking phoneme "makes the left members formally recognizable as bound forms", and these can be coordinated at the surface; this is to be expected under the semantics proposed here, since only the bound morphemes have the right meaning (that is, the right sound) to combine with the meaning of the head.

I have argued in this paper that coordination of parts of words is interpreted at the surface level, through the semantic process of phonological decomposition. No special machinery in needed on top of this, and coordination retains its ordinary meaning. Coordination of parts of words is also subject to prosodic restrictions, which place a minimum size restriction on constituents that can be coordinated. The analysis is superior to deletion proposals, both in its empirical coverage and in its explanation of why parts of words function as independent elements in coordinate structures but not in other grammatical constructions.

References

Ades, Anthony and Mark Steedman (1982). On the order of words. *Linguistics and Philosophy* **4**(4), pp. 517–558.

Artstein, Ron (2001). Conjunction weakening and morphological plurality. In Karine Megerdoomian and Leora Bar-el (eds.), *WCCFL 20 Proceedings*. Somerville, MA: Cascadilla Press.

Artstein, Ron (2002). Focus below the word level. Unpublished manuscript, Rutgers University; presented at CSSP-99 (Colloque de Syntaxe et de Sémantique à Paris), University of Paris 7, October 1999.

Booij, Geert (1985). Coordination reduction in complex words: A case for prosodic phonology. In Harry van der Hulst and Norval Smith (eds.), *Advances in Nonlinear Phonology*, vol. 7 of *Linguistic Models*, pp. 143–160. Dordrecht: Foris.

Booij, Geert and Jerzy Rubach (1987). Postcyclic versus postlexical rules in lexical phonology. *Linguistic Inquiry* **18**(1), pp. 1–44.

Borer, Hagit (1988). On the morphological parallelism between compounds and constructs. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1*, pp. 45–65. Dordrecht: Foris.

Chierchia, Gennaro (1998). Plurality of mass nouns and the notion of "semantic parameter". In Susan Rothstein (ed.), *Events and Grammar*, pp. 53–103. Dordrecht: Kluwer.

Höhle, Tilman N. (1982). Über Komposition und Derivation: zur Konsituentstruktur von Wortbildungsprodukten im Deutschen (On word formation: Constituent structure of compound and derived words in German). *Zeitschrift für Sprachwissenschaft* **1**(1), pp. 76–112.

Kleinhenz, Ursula (1997). *On Words and Phrases in Phonology: A Comparative Study with Focus on German*. Ph.D. thesis, University of Tübingen.

Krifka, Manfred (1990). Boolean and non-Boolean "and". In Lászlo Kálman and Lászlo Polos (eds.), *Papers from the Second Symposium on Logic and Language*, pp. 161–188. Budapest: Akadémiai Kiadó.

Leonard, Henry S. and Nelson Goodman (1940). The calculus of individuals and its uses. *Journal of Symbolic Logic* **5**(2), pp. 45–55.

Link, Godehard (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow (eds.), *Meaning, Use and Interpretation of Language*, pp. 302–323. Berlin: de Gruyter.

McCarthy, John J. (1982). Prosodic structure and expletive infixation. *Language* **58**(3), pp. 574–590.

McCarthy, John J. (1993). A case of surface constraint violation. *Canadian Journal of Linguistics/Revue canadienne de linguistique* **38**(2), pp. 169–195.

Moortgat, Michael (1987). Compositionality and the syntax of words. In Jeroen Groenendijk, Dick de Jongh, and Martin Stokhof (eds.), *Foundations of Pragmatics and Lexical Semantics*, vol. 7 of *Groningen-Amsterdam Studies in Semantics (GRASS)*, pp. 41–62. Dordrecht: Foris.

Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow (1994). Idioms. *Language* **70**(3), pp. 491–538.

Okada, Sadayuki (1999). On the conjoinability of affixal morphemes in English. *Word* **50**(3), pp. 339–363.

Scha, Remko J. H. (1981). Distributive, collective and cumulative quantification. In Jeroen Groenendijk, Theo M. V. Janssen, and Martin Stokhof (eds.), *Formal Methods in the Study of Language, Part 2*, vol. 136 of *Mathematical Centre Tracts*, pp. 483–512. Amsterdam: Mathematisch Centrum.

Smith, George (2000). Word remnants and coordination. In Rolf Thieroff, Matthias Tamrat, Nanna Fuhrhop, and Oliver Teuber (eds.), *Deutsche Grammatik in Theorie und Praxis*, pp. 57–68. Tübingen: Max Niemeyer.

Steedman, Mark (1985). Dependency and coördination in the grammar of Dutch and English. *Language* **61**(3), pp. 523–568.

Steedman, Mark (1987). Combinatory grammars and parasitic gaps. *Natural Language and Linguistic Theory* **5**(3), pp. 403–439.

Wiese, Richard (1992). Prosodic phonology and its role in the processing of written language. In Günther Görz (ed.), *Konvens 92*, pp. 139–148. Berlin: Springer.

Wiese, Richard (1996). *The Phonology of German*. Oxford: Oxford University Press.