Two senses of *complementary distribution* in syntax (teaching): Possible origins, false parallels, and analogies

Carson T. Schütze
UCLA
cschutze@humnet.ucla.edu

0. Recap of Fall '23 AIS Presentation

• phonologists use "complementary distribution" to mean one thing rather clearly: 'where one form (e.g., allophone) occurs, the other(s) never do(es)', which is to say, for any position in any word, (at most) one of the forms (e.g. allophones) can appear there (in a "normal" pronunciation); there are no words such that at a given position, using any of the forms (allophones) would yield a "valid" (i.e. normal) pronunciation

  • I refer to this usage as **CD$_P$**

• syntacticians (as reflected i.a. in textbooks) sometimes use "complementary distribution" to mean the same thing: CD$_P$

  • e.g., "Under the older approach [to Binding], pronouns and anaphors were by definition in **complementary distribution** (pronouns had to be free in their clause, anaphors had to be bound in their clause)" (Carnie 2013: 486). The facts being described include the following:

(1) a. Mary$_i$ likes herself$_i$/*her$_i$.
  b. Mary$_i$'s father likes her$_{i/j}$/*herself$_{i/j}$.
  c. Mary$_i$ believes that she$_{i/j}$/*herself$_{i/j}$ is smart.
  d. Mary$_i$ believes herself$_i$/*her$_i$ to be smart.

  That is, (to a close first approximation) where an anaphor with index i can appear in a sentence, a pronominal with index i cannot, and vice versa.

• those same syntacticians (often in the very same textbooks) also use "complementary distribution" to mean something very different, arguably the exact opposite: I call this **CD$_S$**, and it really means "mutually intersubstitutable", i.e., where one form is possible the other(s) is/are also possible:

  "Possessive pronouns actually behave more like determiners: they are in **complementary distribution** with determiners (***the her** *book*). They appear at the beginning of the noun phrase…*they normally appear* ***in the same syntactic positions*** *as determiners*" (Carnie 2013: 57; emphasis added). …"[*'s*] is in **complementary distribution** with (i.e., cannot co-occur with) determiners (8).…

  8) a) *the building**'s the** roof… [cf. ✓the roof, ✓the building's roof]
  b) *the panther**'s the** coat…
  c) *the man standing over there**'s the** hat…

  … In other words, *'s* and determiners are in **complementary distribution**. As in other domains of linguistics, when two items are in **complementary distribution**, they are instances of the same thing. (Take for example, phonology, where when two phones are

found in different environments—in **complementary distribution**—then they are allophones of the same phoneme.)" [emph added] (Carnie 2013: 209–210)

• moreover, observing CD_S is used to justify a conclusion that is (sometimes explicitly) analogous to the conclusion that phonologists draw from observing CD_P, viz.: the elements that are in "complementary distribution" are versions/flavors/alternate expressions of some single entity, typically a **grammatical category (cf. part of speech)**

• additionally, the nature of the evidence used to justify a generalization of CD_S is typically partly irrelevant and, if applied in general, would lead to wrong conclusions about grammatical category membership; see Carnie's (8) above; another typical paradigm would be

(2) a. **These** books are good.
b. Some books are good.

(3) a. ***These some** books are good.
b. ***Some these** books are good.

Described thus: "{*these, some*} are in complementary distribution: (sentence initially) before a noun you can have one or the other, **but not both**. Therefore, we posit that they are members of the same category (D)."

But logically speaking (2) makes a completely separate point from (1). (1) shows mutual intersubstitutability (in one context). (2) shows that sentences in English happen not to allow the sequence "D D".
If we had instead been trying to characterize the class Adj(ective), this would be obvious:

(4) a. (These/Some) **big** books are good.
b. (These/Some) **old** books are good.

(5) a. (These/Some) **big old** books are good.
b.(?) (These/Some) **old big** books are good.

The distributions summarized in (4) provide evidence that {*big, old*} are members of the same grammatical category. This conclusion is in no way undermined by the goodness of the examples in (5), which simply show that English allows "Adj Adj" sequences[1]

• questions I raised:
1) Why has no one apparently realized that $CD_S \neq CD_P$, when this seems obvious?

2) Would it be pedagogically preferable to avoid the use of one term to mean opposite things (even though no students seem to have been confused by this)?

3) Is there perhaps some level of abstraction at which CD_S and CD_P are two versions of the same thing? (No, but…)

4) If not, does the (sometimes explicit) claim that we're doing syntax "just like phonology" falsely imply that the two fields have more in common methodologically than they actually do, and is that a bad thing?

[1] And indeed, **sounds** that are in CDp **can** occur sequentially: *dungheap* (from a lyric in *Man of La Mancha*) contains a [ŋh] sequence; [ŋ] and [h] are notoriously in complementary distribution in English (but are not allophones, due to lack of phonetic similarity—see below).

• questions raised by the audience:

1) How does this play out in Intro **Ling** textbooks?

2) Does students' apparent lack of (stated) confusion in the face of $CD_P \neq CD_S$ actually reflect a deeper problem? Teaching CD_P is itself notoriously difficult for (many) Intro Ling students;

• several analogies have been brought to bear to help them "get it", but like most analogies they're not perfect—do they actually help or hurt?

• is part of the problem the fact that students don't know what "compleme**n**t(ary)" means (outside linguistics) in the first place? And/or that "complementary" in the expression CD_P might not mean exactly what it means outside linguistics?

1. How prevalent and explicit is this contradictory use of CD_S?

• so far I have found CD_S used in the following **syntax** textbooks (those marked with "*" also use CDp): [2]

Adger (2003)
Beck & Gergel (2014)
Carnie (2013)*
Cowper (1992)
Haegeman (1994)*
Haegeman (2006)
Haegeman & Guéron (1995)
Johnson (2004)*
Koeneman & Zeijlstra (2017)*
Newson et al. (2004)*
Poole (2021)
Sportiche, Koopman & Stabler (2014)*

as well as one research article (Hukari & Levine 1995)

• all Intro **Linguistics** texts I checked never mention CD in the syntax chapter(s):

Fromkin et al. (2000)
Fromkin, Rodman & Hyams (2007, 2017)
Finegan (2012)
Napoli (1996)
Akmajian et al. (2017)
O'Grady et al. (2017)
Roberts (2017)
Language Files 12th edn.

• in assessing the pedagogical consequences, it is interesting to examine exactly what the syntax textbooks say about CD_S:

Beck & Gergel: "Determiners and possessors are said to be in complementary distribution (we can have either one, but not both at the same time). This is typically taken to be evidence that two categories

[2] The inclusion of a particular textbook among those quoted is by no means intended as a criticism. I believe the terminological usage I am highlighting is pervasive; these books are simply reflecting the way syntacticians talk. Being quoted here reflects in part the arbitrary fact that I was able to get access to electronic copies of certain textbooks and automatically search for the phrase *complementary distribution*, and/or this term was indexed. Interestingly, the most prolific syntax textbook author, Radford, never uses CDs, as far as I have been able to determine.

try to **occupy the same slot** in the phrase structure."

• to 'occupy the same slot' is presumably to appear in the same position, the exact opposite of what occurs in CD_P

Adger: "The first important syntactic thing to notice about this set of words [determiners—CTS] is that they are in **complementary distribution**. What this means is that, because they have **exactly the same distribution**, they compete for the same position in the structure of the phrase."

• again, "exactly the same distribution" directly contradicts the definition of CD_P

Sportiche et al (2014: 91):

> "The idea that these are the same category also receives some support from the fact that (in English) only one of these elements can occur at a time, even with verbs like *know* or *say* that allow any one of the three to occur:
>
> (24) You should say { that / if / whether / *if whether / *that if / *whether that } she left
>
> This is what we called complementary distribution in Chapter 2. This fits with the idea that there is a single structural position that can be filled with any one of these elements, but not more than once."

• "single structural position" implies similar distribution, again the opposite of CD_P

BUT elsewhere they use "complementary distribution" to mean CD_P:

> In fact, in these contexts, *she* and *her* are in complementary distribution: where one occurs, the other cannot (and of course we cannot have both at once):
>
> a. She will put a picture of Bill on your desk before tomorrow
> b. *Her will put a picture of Bill on your desk before tomorrow
> c. *She her will put a picture of Bill on your desk before tomorrow
> d. Bill will put a picture of her on your desk before tomorrow
> e. *Bill will put a picture of she on your desk before tomorrow
> f. *Bill will put a picture of she her on your desk before tomorrow
>
> Since they also make the same contribution, picking out a salient female referent, this suggests the conclusion:
>
> (6) *she* and *her* are not only the same category, but are tokens of the same morpheme. Which one can occur depends on syntactic context.

Koeneman & Zeijlstra, who employ both CD_P (for anaphors vs. pronominals, cf. (1)) and CD_S (for determiners vs. possessives) provide the following definition in their glossary:

> "Two elements are in complementary distribution if the presence of one element in a particular position excludes the presence of the other."

• one can get this to apply to both CD_P and CD_S scenarios, but only by taking advantage of an ambiguity:

CD_P: To apply the definition to anaphors vs. pronominals, you have to interpret it to mean that even if you take the first element (say, a pronominal) out of the 'particular position' where it can occur, hence in principle making room for another word, you cannot slot the other element (say, an anaphor) into that (now empty) position:

(6) Mary's father likes <u>her</u>. ⇒ *Mary's father likes <u>herself</u>.

So, here "the presence of one element in a particular position excludes the presence of the other" means 'the fact that one element is attested in a particular position excludes the attestation of the other in that position.'

CD_S: To apply the definition to determiners vs. possessors, you have to take it to mean that the presence of one element excludes the other **if you don't take the first one away**; but if you do take it away, then you CAN put the other element in the same position:

(7) a. ✓<u>Mary's</u> father ⇒ * {<u>Mary's the</u>/<u>the Mary's</u>} father
b. ✓<u>Mary's</u> father ⇒ ✓<u>the</u> father

So, here "the presence of one element in a particular position excludes the presence of the other" means 'when one element is present in a particular position in a particular sentence, this excludes the other element from **also** being present in the same position in the same sentence.'

(For that to even be conceivable, one must construe "position" not as a "slot" that can hold exactly one word, but in some more general way, e.g. in terms of what precedes and what follows. Thus, one could speak of the position in (7) as "between the beginning of the phrase/sentence and *father,*" implicitly allowing that multiple words could in principle appear in that space.

2. New observation that might help explain how CD_S arose

• one can find cases where syntacticians draw the conclusion that some set of forms are members of the same category using data that can be seen as evincing **either CD_P or CD_S**, depending on the "granularity" with which the environments are characterized

• e.g., *for* and *that* are argued to members of the same category (C) based on data like the following:

(8) a. I am excited **that** Mary will be here.
b. I am excited **for** Mary to be here.

• If we characterize the environment as

(9) ... {*excited*/Adj} ___ {*Mary*/DP} Infl VP

then it would be true that the underlined position can host either *that* or *for* (and, I think irrelevantly, not one after the other: **that for*/**for that*): CD_S holds.

• BUT if instead we characterize the environments as

(10) a. ... {*excited*/Adj} ___ {*Mary*/DP} ***will*** VP or a'. ... {*excited*/Adj} ___ {*Mary*/DP} **Infl$_{[+finite]}$**…
b. ... {*excited*/Adj} ___ {*Mary*/DP} ***to*** VP or b'. ... {*excited*/Adj} ___ {*Mary*/DP} **Infl$_{[-finite]}$**…

then CD$_P$ would hold: in environments where *for* is grammatical, *that* is ungrammatical, and vice versa.

• perhaps this is how the use of CD$_S$ arose, i.e. from a failure to notice that genuine cases of CD$_P$ look like cases of CD$_S$ if you squint away the fine detail?

• apropos the question of whether syntax is "just like" phonology, it is worth asking whether observing CD$_P$ ***with no regard to how the environments differ*** (i.e., in the above case only by **sub**categories/features, not by grammatical categories[3]) would justify a conclusion that the forms in CD$_P$ are members of the same grammatical category.

I think NOT.

(11) a. Chris will be **in** {trouble/time/shorts}.
b. Chris will be **very** {tired/unlucky/brave}.

(12) a. *Chris will be **in** {tired/unlucky/brave}.
b. *Chris will be **very** {trouble/time/shorts}.

- so, *in* and *very* are in CD$_P$: where one can appear, the other cannot (of course, one might be able to cook up an environment where both happen to be possible—but let's suppose not).

Obviously no one would conclude from this that *in* and *very* belong to the same syntactic category. This is evidently because the conditioning factor is the **category** of the following element: N vs. Adj, and (at least in English) no one thinks these are subcategories of some single larger category (although they might all be one-place predicates).[4]

- thus, observing CD$_P$ **by itself** does not justify assigning the elements so distributed to the same syntactic category; that evidently requires **intersubstitutability** at some (syntactic) level of generality; observing CD$_P$ is not **sufficient**, and indeed it is not **necessary** either:

---

[3] You may have noticed that "DP" in the environments in (9), (10) cannot stand for identical string sets: I have obscured another subcategorial difference here, viz. DP[NOM] vs. DP[ACC].

[4] Interestingly, Carnie implicitly acknowledges the crucial distinction between categories and subcategories in attempting to run the logic in reverse: he suggests (the idea has probably been around for a long time, but I don't know who first proposed it) that because Adjectives and Adverbs are (allegedly) in complementary distribution (whereby he actually means CDp), they should be viewed as flavors of the same category (which he would like to collectively label "A"). The evidence he adduces looks like this:

(i) a. Chris is **skillful/*quickly**.
b. Chris dances **quickly/*skillful**.
c. a **skillful/*quickly** dancer

However, he fails to remark on cases like the following:

(ii) The scientist made the robot **skillful/quickly**.

So, CDp actually does not hold, if environments are defined as strings. (It could still hold if environments can be stated in terms of structural positions and/or semantic modification relationships.)

e.g.,
these and *those* are (probably?) intersubstitutable everywhere [as long as environments are defined in terms of word strings[5]]; they are certainly **not in CDP**, but they are still (excellent) candidates for membership in the same grammatical category (and subcategory, if Demonstrative is a subcategory of Determiner)

Conclusion: The kind of distributional observation that legitimately leads to the conclusion that two forms belong to the same syntactic category is CD_S, i.e., intersubstitutability; CD_P is not relevant. If CD_S is really the opposite of CD_P, then any claim that syntax and phonology have something in common in this arena is illusory and misleading.

3. Teaching the concept CDp

3.1 How do dictionaries define these things?

complementary:

Merriam-Webster's: "2. serving to fill out or complete; 3: mutually supplying each other's lack"

Collins Concise: "2. "forming a satisfactory or balanced whole"

Oxford Desk: "1. completing…"

complement:

Merriam-Webster's: "1a. something that fills up, completes, or makes perfect"

Oxford Desk: "1. thing that completes"

Collins Concise: "6. *Logic.* the class of all the things that are not members of a given class"

Crystal's (2008: 93) *Dictionary of Linguistics and Phonetics*:

> "**complementary** (*adj*.) (1) A term used primarily in phonology in the phrase **complementary distribution**, referring to the mutual exclusiveness of a pair of sounds in a certain phonetic environment. In English, for example, the voiceless allophone of the /l/ phoneme occurs after initial /p-/, as in *plan*, and the voiced allophone is excluded; conversely, [l] is used initially when no /p-/ precedes. [l̥] and [l] are thus said to be 'in complementary distribution' in this environment. The term is also used analogously in morphology, with reference to the distribution of pairs of forms in grammatical environments (e.g. the selection of alternative forms of plural morpheme in English)."

• when I taught intro ling, I always began this discussion by explicitly noting that *compl**i**ment* and *compl**e**ment* are unrelated, and that the former is not relevant in phonology ([p] does not tell [p^h] that it is looking especially hot today!)

• the notion of completion arguably plays no role in linguistic uses of CD_P, as noted below, although the convention of using "elsewhere" to characterize the most general allophone/allomorph might seem to carry that implication ("elsewhere" does not mean "in every other environment", but rather, "in every other environment where the phoneme occurs at all")

[5] If locations of referents count as part of the "discourse environment" then of course they won't be interchangeably felicitous.

3.2 How do intro linguistics texts explain CD_P? Often using analogies…

Language Files (12th edition) offers the following as their (only) analogy for CD_P:

> Think about what the term complementary means: two complementary parts of something make up a whole. For example, the set of people in your class at any given moment can be divided into the set of people who are under 5'5" tall and the set of people who are 5'5" tall or taller. These two sets of people complement each other. They are mutually exclusive (one person can't simultaneously be both shorter and taller than 5'5"), but **together they make up the whole class**. So also with sounds: sounds that are in complementary distribution do not occur in the same phonetic environments—their distributions complement each other…sound A never occurs in one of B's environments, and B never occurs in one of A's environments (**just like** a person can't be in both the group of people taller than 5'5" and the group of people shorter than 5'5"). (pp. 118–19; emphasis added)

• is this a good analogy? I don't think so. For one thing, "together they make up the whole class" does not carry over transparently to allophonic distributions: this might lead to the expectation that the union of the environments for all of the allophones must constitute the set of all possible phonological environments, but of course that is not true; e.g., /k/ with allophones [k], [kʰ], [kˀ], [kˈ] is phonotactically restricted in English in ways that /t/ is not, e.g. cannot occur finally in a heterorganic coda cluster: *jumped, flicked* vs.* *jumpk, *flitk*

• also, "**just like** a person can't be in both" seems like a false analogy: the reason one person can't be in both groups is because a person has exactly one height, a scalar attribute: one value on a scale cannot fall in two non-overlapping sub-regions of that scale; this is arguably not a contingent fact, but rather, necessarily true given the definition of 'height'

On the other hand, complementary distribution of sounds is clearly a contingent fact, since it can hold in one language but not another (e.g. nasal vs. oral vowels in English vs. French)

• *Language Files* also makes a mess of the point that complementary distribution **by itself** does not entail allophony:

> It's not always possible to find minimal pairs for all distinctive sounds. For example, there are no minimal pairs involving [ŋ] and [h] in English. But it is reasonable to assume that they belong to separate phonemes because they share few phonetic properties, and **no phonological rule determines where they can occur**. (p. 138; emphasis added)

The boldfaced statement is of course false: [h] is restricted to onsets and [ŋ] is restricted to codas in English, so one could trivially write a rule to derive these from a single underlying phoneme; the problem is that these sounds share no properties beyond being consonants (in particular, they differ on the values of [sonorant], [PLACE], [continuant], [voice], [nasal])

Fromkin, Rodman & Hyams (2007):

> "complementary distribution is a fundamental concept of phonology, and interestingly enough, it shows up in everyday life…The first example focuses on *printed* letters…Each printed letter of English has two main variants: lowercase and uppercase…we can formulate a simple rule that does a fair job of determining how letters will be printed: A letter is printed in uppercase if it is the first letter of a sentence; otherwise it is printed in lowercase…the two variants of a single mental representation (phoneme or letter) are in *complementary distribution* because they never appear in the same environment. And, substituting one for the other…may sound or look unusual, but it will not change the meaning of

what is spoken or written." (p. 265)
[their second analogy involves letter forms in cursive writing]

Added for the 10th edition:

> "Finally, Superman and Clark Kent, Batman and Bruce Wayne, and Dr. Jekyll and Mr. Hyde…are in complementary distribution *with respect to time*. At a given moment in time, the individual is either one or another of his alter egos." (p .227)

Pam Munro (p.c.)

• shoes (CD) vs. socks (free variation)

3.3 Evaluating the analogies for CD_P

Possible desiderata	**Upper- vs. lower case letters**	**Superhero alter-egos**	**Shoes**
Allows for a way to compare CD with free variation	✓? (one might try "ɡ" vs. "g")	✗?	✓
Captures the (usually) nonarbitrary connection between forms and environments	✗? (tHIS IS LOGICALLY POSSIBLE.)	✓ we find Wonder Woman and not Diana Prince in "environments" characterized by someone being in danger "because" WW has (super-)powers that are useful for rescuing people from danger, while Diana doesn't[6]	✓ (a right shoe is shaped like a right foot—putting it on a left foot would be physically challenging)
Captures the phonetic similarity criterion for allophony?	✗? ("ɑ" is arguably more similar to "D" than to "A"; relationship between "g" and "G" seems arbitrary)	✓ (alter egos have the same height, weight, eye color, etc.)	✓? (while a pair of shoes being the same size reflects the symmetry of feet, the fact that they are the same color etc. is perhaps an arbitrary social convention)

[6] Superheroes may vary in the extent to which their "superpowers" are tied to their superhero guise, at least insofar as these rely upon the accoutrements of that guise. E.g., Wonder Woman can deflect bullets using her magic bracelets, which (somehow) become available as part of her transformation (spinning around, triggering an "explosion")—Diana Prince presumably doesn't have those bracelets in her jewelry box. But Superman's powers aren't tied to the cape and tights or stepping into a phone booth: Clark Kent could use X-ray vision if he felt like it.

References

Adger, David (2003). *Core syntax: A minimalist approach.* Oxford: Oxford University Press.

Akmajian, Adrian, Ann K. Farmer, Lee Bickmore, Richard A. Demers & Robert M. Harnish (2017). *Linguistics: An introduction to language and communication*, 7th edn. Cambridge, MA: MIT Press.

Beck, Sigrid & Remus Gergel (2014). *Contrasting English and German grammar: An introduction to syntax and semantics.*

Bromberger, Sylvain & Morris Halle (1989). Why phonology is different. *Linguistic Inquiry* 20, 51–70.

Carnie, Andrew (2013). *Syntax: A generative introduction,* 3rd edition. Malden, MA: Wiley-Blackwell.

Cowper, Elizabeth A. (1992). *A concise introduction to syntactic theory: The Government-Binding approach*. Chicago: University of Chicago Press.

Crystal, David (2008). *A dictionary of linguistics and phonetics*, 6th edition. Malden, MA: Blackwell.

Dawson, Hope C. & Michael Phelan (eds.) (2016). *Language files: Materials for an introduction to language and linguistics, 12th edition*. Columbus: The Ohio State University Press.

Finegan, Edward (2012). *Language: Its structure and use,* 6th edn. Boston: Wadsworth.

Fromkin, Victoria, Robert Rodman & Nina Hyams (2007). *An introduction to language,* 8th edn. Boston: Thomson Wadsworth.

Fromkin, Victoria, Robert Rodman & Nina Hyams (2017). *An introduction to language,* 11th edn. Boston: Cengage.

Haegeman, Liliane (1994). *Introduction to government and binding theory,* 2nd edition. Oxford: Blackwell.

Haegeman, Liliane (2006). *Thinking syntactically: A guide to argumentation and analysis*. Malden, MA: Blackwell.

Haegeman, Liliane & Jacqueline Guéron (1999). *English grammar: A generative perspective.* Oxford: Blackwell.

Hukari, Thomas E. & Robert D. Levine (1995). Adjunct extraction. *Journal of Linguistics* 31, 195–226.

Johnson, Kyle (2004). *Introduction to transformational grammar*. Ms., UMass Amherst.

Koeneman, Olaf & Hedde Zeijlstra (2017). *Introducing syntax.* Cambridge: Cambridge University Press.

Marantz, Alec (1991). Case and licensing. In Germán F. Westphal, Benjamin Ao & Hee-Rahk Chae (eds.), *ESCOL '91: Proceedings of the Eighth Eastern States Conference on Linguistics*, 234–253.

Napoli, Donna Jo (1996). *Linguistics*: *An introduction.* New York: Oxford University Press.

Newson, Mark, Marianna Hordós, Dániel Pap, Krisztina Szécsényi, Gabriella Tóth & Veronika Vincze (2004). *Basic English syntax with exercises*. http://primus.arts.u-szeged.hu/bese

O'Grady, W., J. Archibald, M. Aronoff & J. Rees-Miller (eds.) (2017). *Contemporary linguistics: An introduction. Seventh edition*. Bedford/St. Martin's.

Poole, Ethan (2021). Ling 200B lecture notes, UCLA.

Roberts, Ian (2017). *The wonders of language; or how to make noises and influence people.* Cambridge: Cambridge University Press.

Sag, Ivan A., Thomas Wasow & Emily M. Bender (2003). *Syntactic theory: A formal introduction,* 2nd edn. CSLI.

Sportiche, Dominique, Hilda Koopman & Edward Stabler (2014). *An introduction to syntactic analysis and theory*. Wiley-Blackwell.