

Fiat Lingua

Title: Potential Paonese: A Reconstruction from Jack Vance's *The Languages of Pao*

Author: Logan Kearsley

MS Date: 10-14-2015

FL Date: 11-01-2015

FL Number: FL-000032-00

Citation: Kearsley, Logan. 2015. "Potential Paonese: A Reconstruction from Jack Vance's *The Languages of Pao*." FL-000032-00, *Fiat Lingua*, <<http://fiatlingua.org>>. Web. 01 November 2015.

Copyright: © 2015 Logan Kearsley. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.



<http://creativecommons.org/licenses/by-nc-nd/3.0/>

Potential Paonese

A Reconstruction from Jack Vance's *The Languages of Pao*

The Languages of Pao, by Jack Vance, explores the strong Sapir-Whorf hypothesis in the description of a stagnant people, the Paonese, of the planet Pao. The initially uniform Paonese culture is fractured by a massive social engineering project whose central component is the introduction of numerous artificial languages, each designed to produce in its speakers a tendency towards a different socially useful pursuit. Very little description is given to any of these languages, however, with the exception of the original Paonese.

Even the original Paonese, however, is not a real language. That may seem obvious given that it is a fictional language spoken by a fictional people, but well-developed conlangs like Quenya, Klingon, Dothraki, Na'vi, and numerous less well-known examples can reasonably be called “real languages” in the sense that, no matter the real or fictional status of their native speakers, they exist in the real world; one can learn them, compose novel sentences in them, and generally use them as a practical medium of communication. None of the languages of Pao meet that standard; they are strictly imaginary. They exist only as ideas of languages from a fictional world.

Of the original Paonese, however, we do have scattered words and descriptions by the narrator, who represents a contemporary, though linguistically naïve, observer from the fictional world. In the attempt to understand the languages of the fictional world of Pao, this puts us in a similar position as philologists attempting to piece together the scattered fragments of a poorly attested ancient language. There is too little known to provide a definitive description or reconstruction. However, we may still endeavor to construct a real language that *could have been* Paonese.

Purpose of this Article

The purpose of this article is not to provide a complete linguistic description, reference grammar, or lexicon of the reconstructed Paonese language. Rather, my intent is to document the evidence for Paonese and the reasoning behind the decisions that I made in producing a reconstruction. To that end, while this account is significantly cleaned up compared to my original stream-of-consciousness notes, there is still a good deal of meandering- presenting a possibility, noting its problems, presenting possible alternatives, noting their problems, and sometimes ending up back at the beginning again. It is my hope that, rather than being seen as a pointless diversion, the reader will find these wanderings down blind alleys, investigating solutions that end up getting thrown out, to be a valuable source of insight into the reconstructive process, odd back-alleys of English pronunciation, and the reasons why certain final decisions were made.

It is particularly important to note that the resulting “potential Paonese” is not necessarily *the* Paonese. It most likely *does not* represent the original Paonese as it existed in the mind of Jack Vance; and indeed, in the strictest sense that may be impossible, as I am unconvinced that that original Paonese language ever actually existed at all beyond the scraps presented in the book. Other equally valid reconstructions, based on different interpretations of the evidence and different arbitrary decisions about unattested features are certainly possible.

Additionally, the purpose of this specific reconstruction is not to produce the simplest

possible language that accords with Vance's work, with the simplest possible explanations for all evidence. Rather, it is to develop a language which is naturalistically complex in ways that are not obviously derived from a monoglot Anglophone perspective, and yet still *can* explain the evidence in the book.

Plot Summary

At the beginning of the novel, the planet Pao is a mostly agrarian backwater world with a culturally homogenous population ruled by an absolute monarch, known as the Panarch. Knowing that their lack of diversity makes Pao vulnerable to external economic and military powers, the Panarch attempts to hire a consultant from the planet Breakness, Lord Palafox, to plan a reformation of their society. The Panarch, however, is betrayed, assassinated by his brother, who seizes the throne, and the rightful heir, Beran, is kidnapped and taken to be raised on Breakness as a political hostage.

After Pao is conquered and made to pay tribute to the planet Batmarsh, the new Panarch again seeks aid from Lord Palafox, who has a plan to create castes of warriors, scientists, and merchants by segregating Paonese society via the forcible introduction of new, mutually-unintelligible languages, to be taught to selected populations of Paonese children. Accomplishing this goal requires the training of a corps of language instructors from Breakness, fluent in the new constructed languages and capable of teaching them to the Paonese. Beran infiltrates this corps in order to return to Pao incognito, and participates in the creation of an additional pidgin called Pastiche- a mixture of all of the constructed languages of Pao used among the instructors.

Eventually, Beran reveals himself to the people of Pao, and support for the usurping Panarch crumbles. The warrior caste is able to repel their invaders, and Beran reclaims his title, and the planet enters a brief golden age of wealth and progress. Displeased with the divisions caused by Palafox's language program, however, Beran eventually attempts to re-integrate the castes. This results in an uprising and coup staged by the warriors. Beran convinces them, however, that they will be unable to rule the planet alone as they share no common language with the rest of the populace; they therefore allow him to remain in office, and agree to his decree that every Paonese child must learn Pastiche, once again uniting the planet under a common language and in a common culture- but now possessing the diverse linguistic attributes of proud warriors, skilled scientists, and savvy merchants.

Overview of Attested Paonese

From what limited explicit description is given, we know that Paonese is polysynthetic, contains no verbs or adjectives, and is composed mainly of “nouns, post-positions, and temporal indices”. It is said to be derived from another language known as “Waydalic”, but “molded into peculiar forms”; no further information is given about Waydalic. Paonese is also said to have no formal comparison (which we might infer to mean “grammatical comparison”), no words for “good”, “better”, or “best”, and no words for "prestige", "integrity", "individuality", "honor", or "justice". In more impressionistic terms, Paonese is said to “[present] a picture of a situation rather than describing an act”.

The lack of formal comparison combined with the absence of lexemes for “good”, “better”, or “best” potentially has interesting Whorfian implications, and suggests that a reconstructor should attempt to avoid introducing any means for a Paonese speaker to express overt relative evaluation or feelings of superiority or inferiority between different situations or

people¹. In a living language, it would be nearly impossible to prevent people from finding some way of doing so if they so desired, but it says something about Paonese culture that they apparently do not so desire- they do not “feel” the lack. A parallel can be drawn here to the linguistic culture of the Anarresti people from Ursula K. Le Guin’s novel *The Dispossessed* (1974; for comparison, *The Languages of Pao* was published in 1958). While practically nothing of their language is actually known, apart from proper names, it is known that the Anarresti culturally suppress the usage of possessive pronouns and egocentric expressions, even though the mechanisms for such are technically available to them². Something similar may be going on in Paonese, where comparative constructions have not been grammaticalized due to strong social disapproval of those who attempt to express them.

The remaining lexical requirements, while also meant to be indicative of something interesting about Paonese culture, are, however, of least use to the reconstructive conlanger, since it is a trivial matter to simply omit certain lexical items from the dictionary.

Only a single complete sentence of Paonese is ever given. However, there are numerous titles and names of people and places given throughout the book which may be able to give us some information on Paonese phonology and morphology. None are ever explicitly identified as belonging to any particular language, but we can make inferences based on the provenance of characters and the names of places- i.e., a native of Pao is likely to have a genuinely Paonese name, while names of foreigners are probably not good evidence of Paonese. Similarly, cities built for and inhabited by speakers of other languages, despite being located on Pao, may not be good Paonese names, while pre-upheaval names of Paonese landforms and cities inhabited by traditional Paonese, with a few obvious exceptions, almost certainly are.

Several factors point to the aforementioned conclusion that the narrator is a linguistically naïve observer: First, the narrator is unclear about the distinction between suffixes and postpositions; this could be a matter of genuine uncertainty about the proper analysis, or failure to clearly explain what the narrator himself actually understands well, but is equally well explained by a naïve narrator who simply doesn't know the difference. Second, the narrator makes use of non-standard terminology in cases that do not present any obvious improvement for understanding by a lay audience, such as using "volition" in place of "agent", and “agency” where the more standard term would be “instrument”. Finally, it is known extra-fictionally that the author, Jack Vance, had little to no formal linguistic training³, and thus cannot be expected to have produced a narrator who is himself savvy in linguistic analysis. He did, however, travel extensively, and thus would have been exposed to numerous foreign names and spelling conventions which may have been reflected in his work (Williams 2006). With these observational justifications, the assumption of a linguistically naïve narrator provides significantly increased latitude in interpretation for the reconstructive conlanger. This is both a blessing, in that it allows us to resolve some apparent problems by arguing that the narrator's descriptions were simply wrong, but also a curse, in that we lose confidence in some of the evidence we have and thus must put more effort into original creation.

1 Thanks to John Quijada for pointing this out.

2 Intrafictionally, the Anarresti language, Pravic, is actually a conlang, voluntarily adopted by the original settlers of Annares in order to further separate them from the culture of their homeworld, Urras, from which they were political refugees.

3 Vance enrolled in an Army Intelligence program to learn Japanese, but washed out (Williams 2006); this may explain where he got the idea for postpositions in Paonese, but I could find no documented evidence that he seriously studied any other language or linguistics in general.

Attested Words

There are three Paonese titles: “Panarch”, “Medallion”, and “Ayudor”. The first two are fairly obviously either borrowings or translations for the benefit of the reader (“Panarch” from Greek roots meaning “ruler of all”, “Medallion” from English). This leaves only “Ayudor” as an example of genuine Paonese, apparently meaning something like “regent”.

There are ten known Paonese personal names: “Aiello Panasper”, “Bustamonte”, “Beran Panasper”, “Vilnis Therobon”, “Mornune”, “Hessenden Andrade”, “Est Coelho”, “Gitan Netsko”, “Ercole Paraio”, and “Can”. Paonese names in the book follow the “given name” – “family name” pattern, but it is unknown whether that is indicative of how names are actually used by the Paonese, or another adaptation for the reader. Several of these names, however, appear to be derived from real names on Earth, and might not be considered good evidence of Paonese. “Panasper” could easily be a compound of Greek “pan” and Latin “asper”⁴; “Aiello” is an Italian name (typically a surname, but conversion of surnames to given names is not unusual); “Andrade” and “Coelho” are both Portuguese surnames; “Bustamonte” is a minor variant of “Bustamante”, a surname originating in Spain; and “Ercole” is the Italian equivalent of Hercules.⁵ This leaves us with “Beran”, “Can”, “Est”, “Gitan”, “Hessenden”, “Mornune”, “Netsko”, “Paraio”, “Therobon”, and “Vilnis” as probable good evidence for Paonese.

There are five unambiguous names of Paonese cities: “Eiljanre”, “Koroi”, “Sherifte”, “Donaspara”, and “Spyrianthe”. In addition to the cities, we have 24 other assorted place names: “Pergolai”, “Jhelianse Sea”, “Hylanth Littoral”, “Hylanthus Sea”, “Hyaline Gulf”, “Zelambre Bay”, “Mervan Pond”, “Feraï”, “Viamne”, “Cantatrino”, “Moravi Inn”, “Muniment Library”, “Plarth”, “Qurái Peninsula”, “Fraevarth”, “Sgolaph Mountains”, “Mount Droghead”, “Lido”, “Mathiole”, “Pamalisten”, “Vredeltope”, “Maesthgelai Peninsula”, “Rovenone Canal”, and “Beauclare Quarter”.

“Mount Droghead” and “Beauclare Quarter” are a bit problematic; it is difficult to imagine “head” being anything other than the homographous English word, leading us to wonder what a “Drog” might be, and “Beauclare” is a real French surname. Perhaps this is an indication of some French contribution to the original settlement of Pao, in addition to the Italian and Portuguese influences, but it can hardly be considered a normal Paonese word.

We also have the names of Pao's eight continents, whose names are transparently derived from the basic numbers of Paonese's base-eight numeral system: Aimand (Ai), Shraimand (Shrai), Vidamand (Vida), Minamand (Mina), Nonamand (Nona), Dronamand (Drona), Hivand (Hivan), and Impland (Imple).

From the one attested complete sentence, we get the words “Rhomel”, “bogal”, “mous”, and “es”, along with various other bits of graphological material which seem to represent exclusively bound morphemes rather than independent words.

Finally, there are a few additional miscellaneous words of probable Paonese origin: “Mamarone”, an elite soldier; “praesens”, a “vitality word” of unclear specific meaning; “Auriol”, Pao's sun; “Kanetsides”, a holiday; “furze”, possibly a kind of wool (collected by “furze-cutters” who are apparently comparable to shepherds); and, of course, “Pao” itself.

4 Rough, unrefined, sharp, or newly minted; from <https://en.wiktionary.org/wiki/asper>

5 As previously mentioned, Jack Vance did a great deal of world travelling, including a stay in the Italian village of Positano, in which he wrote *Vandals in the Void* in 1951 (Williams 2006), and on which he based a later novel, *Strange People, Queer Notions*. It is thus unsurprising that we should see several examples of specifically Italian and other Romance names dusted off for use in *The Languages of Pao*.

Analysis and Reconstruction

Phonology

No attempt is ever made in the book to explain the proper pronunciation of any Paonese words. We are left to guess at what the pronunciations may be based on the assumption that they represent only the attempt of a native English speaker to transcribe foreign words according to his own spelling system, with no particular thought given to consistency in transliteration. Unfortunately, English spelling is famous for its inconsistency, and we are left with no choice but to simply guess at many of these words. Doubly unfortunately, it is impossible to tell for sure if Paonese may have, or been intended to have, phonemes entirely foreign to English; this reconstruction will therefore necessarily produce a phonetic inventory that is very similar to that of English, despite attempts to inject plausible variations from Anglophone phonology where the evidence is ambiguous. Certain words, however have unusual features that may provide more objective evidence for some non-Anglophone aspects of Paonese pronunciation.

The place name “Sgolaph” is unusual in its apparent juxtaposition of an unvoiced [s] and voiced [g]. Although voicing of the <th> digraph is uncertain, a similar situation occurs in the place name “Maesthgelai”. In the latter case, this could be explained by an apparent cluster actually being broken across a syllable boundary; in the first case, however, the word-initial environment means that <sg> cannot be anything but a single syllable-onset cluster. Of course, English frequently uses the grapheme <s> for a voiced /z/ sound as well as /s/, and Italian specifically employs the spelling <sg> for [zg] in initial positions, as in *sgridare* (“to scold”)⁶. We have already identified some Italian influence, so this may have been the way Jack Vance intentionally chose to spell [zg].

There are several ways to approach “Sgolaph”, but let’s first back up and discuss the matter of voicing in general. English has a phonemic voicing distinction in all of our stops and most fricatives (with the exception of /h/). To a native English ear, this seems pretty straightforward- “b” is voiced, “p” is not; “d” is voiced, “t” is not; and so forth. At a phonetic level, however, things are not so simple, and this can result in Anglophone listeners mis-identifying the relevant phonemic distinctions in languages that do not have the same range of allophonic variation as English, and distinguish different series of stops and fricatives by different particular combinations of phonetic features. So, while it is tempting to simply declare that the phonemes of Paonese correspond exactly to how an English speaker would pronounce the Paonese words as Jack Vance wrote them⁷, there are good reasons to dig a little deeper. Stopping there would result in a fairly *boring* phonology; and furthermore, it is simply *unlikely* that a far-future language of an alien⁸ culture like Paonese would happen to reproduce English phonology so closely. We can do better.

The English stops /b/, /d/, and /g/ are only fully voiced in intervocalic positions, and are only partially voiced in initial and final positions (i.e., following or preceding silence).

⁶ Thanks to John Quijada for the suggested example.

⁷ It is also tempting to say that Paonese phonemes correspond exactly to the *letters* of Paonese words as Jack Vance wrote them; but the fact is that Jack Vance was an English speaker, writing for English speakers, most likely using the conventions of English orthography to convey his intended sounds. And given the weirdness of *English* orthography, it is a safe assumption that neither he nor any of his readers would ever have used such simplistic pronunciations.

⁸ “Alien” in the sense of “not from around here and unfamiliar to us”, not in the sense of “non-human beings from outer space”.

Accordingly, many English phoneticians prefer the categorization of fortis (stronger airflow) vs. lenis (weaker airflow) to distinguish /p/, /t/, /k/ from /b/, /d/, /g/, respectively (Yavaş 2011). In particular, when preceded by an /s/, even across syllable or word boundaries boundaries, the phonetic distinction between “voiced” and “unvoiced” English stops is purely lenis vs. fortis; there is no difference in actual phonetic *voicing* in minimal pairs such as “discussed” (/diskʌst/) vs. “disgust” (/disgʌst/) or “disperse” (/dispə-s/) vs. “disburse” (/disbər/).

Given the relative weakness of the phonetic voicing distinction in English, two other major features serve to help distinguish the fortis and lenis stops. First, vowels and nasals are lengthened before lenis (voiced) stops relative to their pronunciations before fortis stops. Second, lenis (voiceless) stops are aspirated in stressed syllable-initial positions. It turns out that simply dropping the aspiration from a voiceless stop in a stressed position is sufficient for English speakers to confuse it with a lenis (voiced) stop; the onset of voicing in a following vowel is sufficient to move such stops into the perceptual range of allophones of the partially-voiced lenis stop series. This can cause a great deal of difficulty for English speakers trying work with languages that have a phonemic aspiration distinction, a fact with which I became intimately familiar over a semester of trying to elicit linguistic data from a speaker of Eastern Armenian

If Paonese distinguishes its stop series by aspiration, this would provide us with some flexibility in re-analyzing the voicing status of various stops in the attested corpus, but, more importantly, would allow us to make strong inferences about where stresses must fall in order to produce perceptions in an Anglophone listener of voiced and voiceless stops in the attested positions.

Eastern Armenian makes a three-way distinction between fully voiced stops, voiceless unaspirated stops, and aspirated stops. Hindi goes even further, with four series of stops covering all combinations of voiced vs. voiceless and aspirated vs. unaspirated. There are, however, good reasons for deciding on a two-way distinction based purely on aspiration in Paonese: mainly, the fact that assuming only a two-way distinction with no phonemically voiced stops provides much greater constraints on our potential guesses on other features of proper pronunciation. If we assume that they are actually all simply unaspirated, we avoid having to make separate guesses as to the phonetic status of every instance of an apparently-voiced stop. Additionally, we can conclude that apparently-voiced stops cannot occur in unstressed syllables; otherwise, they would have been properly perceived as unvoiced by the Anglophone narrator, and transcribed as such (with a few regular exceptions; see below). This allows us to determine the stress patterns of several words. Unfortunately, we are left guessing which unvoiced stops are aspirated or not in unstressed syllables.

There are four words with absolutely unambiguous cases of apparently-voiced stops occurring as the entirety of a syllable onset due to being the initial consonant of a word:

Beran Donaspara Gitan bogal

The first three of these must have initial stress. The last, “bogal”, however, presents a potential problem with two apparently-voiced stops attempting to attract the stress. If we make the conservative syllabification assumption that a single intervocalic consonant will be grouped as the onset of the following syllable, rather than as the coda of the preceding syllable, then we also have four more possible examples:

Ayudor Lido Vredeltope Therobon

Depending on the placement of stress in these first three words, the <d> in each case may in fact be in a position where /d/ and /t/ are neutralized to /r/ in English; thus there is more than one reason why a /d/ may have been misheard here, and we cannot conclude that it *must* be because of following stress. Additionally, it is possible that, just as aspiration is an allophonic variation in English unvoiced stops, voicing may be an allophonic variation in Paonese unaspirated stops. All four of these words, plus “bogal” from the first set, have apparently-voiced stops in intervocalic positions, a prime environment for triggering allophonic voicing. This would producing voiced stops that are indistinguishable from other unaspirated stops to a native Paonese speaker, but are heard, and thus written, differently by the Anglophone narrator. Since only the second voiced stop in “bogal” can be explained by allophonic voicing, we can conclude that the stress in that word must again fall on the initial syllable after all.

Three more words are ambiguous due to uncertainty about proper syllabification:

Hessenden Maesthgelai Pergolai

Under the assumption that a syllable boundary falls between <n> and <d>, “Hessenden” has final stress, unless nasals also contribute to the environment for allophonic voicing.

Under the assumption that a syllable boundary falls between <th> and <g> and between <r> and <g>, “Maesthgelai” and “Pergolai” both have penultimate stress⁹- unless /r/ also contributes to the environment for allophonic voicing.

If we ignore the complications of neutralization and possible intervocalic voicing, we end up with three examples of unambiguous initial stress, four of final stress, three examples of some form of medial stress, and one indeterminate, with no obvious patterns to determine the rules for each. The simplest conclusion in that case is that Paonese has unpredictable lexical stress. That's a rather unsatisfying result, as it tells us nothing useful about the rest of the attested corpus. If, however, we allow for allophonic voicing of unaspirated stops, then we must throw out several ambiguous examples, but we are left with four examples of initial stress and one definite and one possible example of medial stress, where the possibly-medially-stressed words (“Pergolai” and “Maesthgelai”) appear to share a suffix <-lai>, which provides a reasonable morphophonological explanation for why their stress might shift.

To motivate the identification of this repeated sequence as a morphological suffix, rather than a mere phonological coincidence, we can even make a fairly good guess at what it means. No other words end in <lai>, but four other words in the corpus end in <ai>: Ai, Shrai, Ferai, and Qurai. “Ai” and “Shrai” are numerals, and presumably monomorphemic, but “Qurai” is the name of a peninsula, as is “Maesthgelai”, and “Ferai” is the name of an island. It is not unusual for natural languages to use a single word for “island” and “peninsula”, and so the common string-suffix may actually be a genuine morphological suffix (or compounded root) meaning “island/peninsula”. In this case, the actual suffix may be <ai>, with “Pergolai” and “Maesthgelai” only incidentally sharing an <l>, or some morphophonological rule may have deleted the <l> from a <-lai> suffix in the case of “Qurai” and “Ferai”. In the interest of explaining more data all at once, we will assume a basic form of <-lai> for the suffix, and an *l*-deletion rule.

⁹ Or antipenultimate, if <ai> is to be interpreted as two syllables, rather than as a diphthong. My intuition, however, is that Vance would have been likely to use an apostrophe or some similar convention to separate the vowels if they were meant to be separate syllables, and it is most reasonable to assume that terminal <ai> does indeed represent a diphthong like /ai/.

Now, we can return to the matter of “Sgolaph”. This word should have initial stress. Unvoiced stops in English are *not* actually aspirated after /s/, but they are more fortis than phonemically-voiced stops. It is not obvious just how much fortition a native Paonese speaker would put into a normal unaspirated stop, so this could really go either way; a sufficiently lenis [k] would be mis-heard as /g/.

The known Italian influence on Pao suggests that <sg> should actually be [zg], use <s> as /z/ intentionally, and this is plausible if a voiced fricative contributes to allophonic voicing just like nasals and vowels and /r/ might. However, while there are several instances of <s>+stop (in “Est” and, of course, “Sgolaph”, where the cluster must be contained in a single syllable; and in “Netsko” and “Donaspara”, where it may be across syllable boundaries) and an <f>+stop (“Sherifte”, which may be also be across a syllable boundary), there are no obvious instances of any clusters of a voiced fricative+stop, which suggests that *this* <s> should also still be treated as an /s/, not a /z/. But that evidence is also not conclusive; it is entirely possible that clusters with voiced fricatives are allowed, and just rare. In English, after all, we restrict voiced clusters /zd/ to coda positions (as in “buzzed”¹⁰) or split across syllable boundaries, while all three unvoiced /s/ +stop clusters (/st/, /sp/, /sk/) can occur anywhere, but they *do* exist. Furthermore, while the environments are not identical and thus there *may* be variation in lenis/fortis pronunciations, a word like “Netsko” should also have come out as “Netsgo”, if we assume initial stress (so the <k> is not necessarily aspirated) and uniform fortition- unless the <k> is aspirated. The same argument applies to “Spyrianthe”- the environment is the same, so either the <p> is aspirated, the <g> really is voiced, or there is a difference in the fortition of Paonese /k/ vs. /p/.

With no extrafictional evidence, I would still be more inclined towards the “lenis pronunciation” = “misheard /sk/” explanation, because that results in a more restricted, and thus more distinctive, phonology. However, since we know that Vance was exposed to Italian, and borrowed Italian names, I think we must come down on the side of the fence that admits that [zg] is the actual intended pronunciation. We thus conclude that *any* voiced segment, including fricatives, can trigger allophonic voicing in unaspirated stops, and that Paonese actually allows clusters of phonemic /zk/ (realized as phonetic [zg], and presumably /zt/ and /zp/ as well) at low frequency.

The typical level of fortition in Paonese unaspirated stops, while now irrelevant to the interpretation of “Sgolaph”, does constrain the interpretation of every other instance of <s>+stop as aspirated or unaspirated; if the average fortition level is low, they must all be aspirated. If it is high, aspiration is ambiguous. Low fortition thus provides greater constraints, which is attractive- but it results in a corpus that contains *no* instances of <s> followed by an unaspirated stop, without providing any evidence that such sequences *ought* to be disallowed. That just feels wrong, so I will assume a fairly uniform level of high fortition in all stops (or at least those following /s/), so as to be able to introduce a better mixture of aspirated and unaspirated stops into the corpus arbitrarily.

As a final note on voicing, we should address what happens to unaspirated stops on word boundaries, where they may have a voiced segment on one side but silence on the other. For initial stops, we could assume that they become partially voiced, in anticipation of the following vowel; since Paonese does not distinguish voicing, there is no pressure for speakers *not* to make that “error” to ease pronunciation, and eventually cement it as an allophonic rule. But, remember that English voiced stops are themselves only *partially* voiced in initial positions! Thus, such a rule in Paonese would make all initial unaspirated stops sound exactly like English voiced stops

¹⁰ Thanks to Jim Henry for the suggested example.

in all conditions, removing our justification for identifying stress. In order to preserve the integrity of our analysis so far, we *must* therefore assume that this variation does not occur in standard Paonese pronunciation, and unaspirated stops which are preceded by silent or another unvoiced segment *do* remain completely devoiced.

Allophonic voicing rules for the end of a word, however, must account for the proper perception of the names of continents, which all end in <-mand>. English /nt/ in terminal position is not aspirated, and English speakers can still reliably tell the difference between [nt] and [nd]. Differences in vowel length could account for that, but if we assume that Paonese speakers do not change their vowel lengths in response to the voicing of following consonants (and why should they? That is an English feature, not a universal human one!), that should tilt Anglophone perception towards an accurate /nt/ instead. The simplest explanation for this appears to be that a preceding nasal will cause voicing in a following unaspirated stop in all positions, even if it is word-terminal and not sandwiched with other voiced sound. This implies that the terminal <t> in “Muniment” must be aspirated- otherwise, it would have undergone voicing. The only other example we have a pure-Paonese word ending in a stop is “Vredeltope”¹¹ (assuming a silent <e>); in that case, either the <p> is aspirated, or a vowel, unlike a nasal, does not trigger allophonic voicing, or both. In order to again allow more freedom in assigning aspirations, we will assume that only nasals trigger voicing in stops that are not followed by another voiced segment (including at the end of a word). There is some theoretical justification for this restriction, as homorganic nasals and plosives can be perceived as a single unit (a prenasalized stop), where a vowel and following consonant would not be, Terminal unaspirated stops may also be unreleased, although this is not strictly necessary to explain the differences in Anglophone perception.

We can thus formulate a few general rules for Paonese phonology so far:

1. Paonese roots have default initial stress.
2. Stress may be modified by stress-attracting affixes, such as *-lai*.
3. Unaspirated stops undergo voicing between other voiced phonemes or after a nasal.
4. /l/ is deleted after /r/.

“Qurai”, “Jhelianse”, and “Rhomel” provide some evidence for the possibility of additional non-English sounds in Paonese by the simple fact that they appear to violate English orthographic rules. It would seem an odd coincidence if “Qu” were not intended as a digraph, as it is in English. However, that digraph, in languages that use it, is typically restricted to representing /k/, /kw/, or /k^w/, while other letters are already widely attested for /k/ and there is no other evidence of a Paonese /w/. Thus, we will assume the <u> in this case actually represents an independent vowel, and <q> is meant to represent some other sound. Somewhat arbitrarily, we will assign <q> its IPA value as a voiceless uvular stop.

The apparent digraph <jh> is more problematic. In Walloon, it represents a phoneme which varies in realization between /h/ and /ʒ/ according to dialect. In Esperanto and the romanization of Kurdish¹², it’s also a notation for /ʒ/. If, however, <jh> means /ʒ/, then what is the value of plain <j> in “Eiljanre”? It could represent its standard English value of /dʒ/, but if so then that is the only attested instance of a voiced affricate, and one of only three possible

11 We also have a marginal example in “Drog” (from “Droghead”), but I’m hesitant to use that as evidence for or against any feature of word-final stops in Paonese since the only place it appears is compounded with English.

12 Although the amount of exposure Jack Vance may have had to Esperanto or romanized Kurdish is questionable, there isn’t a whole lot else to go on.

affricates of any kind. The other potential examples are both /ts/, but in both cases (“Netsko” and “Kanetsides”) it is equally reasonable to assume that there is an intervening syllable break and <ts> does not actually represent an affricate. The most straightforward solution seems to be to assume that <j> always represents /ʒ/, and <jh> is actually an accurate representation of /ʒh/, but a cluster of voiced and voiceless fricatives at the beginning of a word would be truly odd.

Another option is to assume that Vance added an ‘h’ to the spelling simply so that readers would be less likely to pronounce “Jheliense” with an initial /dʒ/ (the typical English value for <j>)- in which case, perhaps it is simply an orthographic convention for word-initial /ʒ/, of which we happen to have only a single attested example. The medial <j> in “Eiljanre” is then still pronounced simply as /ʒ/, and written differently strictly as a matter of orthographic convention. If the <jh> digraph appeared in any other positions, or if there were any other initial <j>s with which it could contrast, there would be a stronger case for analysis as a weird non-English phoneme of some sort; but with no other obvious leads to go on, this seems like the best conclusion.

Handling “Rhomel” seems relatively straightforward. The initial <rh> *could* be interpreted as an odd, but apparently fine for Paonese, sequence of two phonemes, /rh/. However, the digraph <rh> is also used fairly consistently across different languages in the Roman alphabet to indicate what was at one point a voiceless ‘r’ sound, as in “rhapsody”, from Classical Greek ῥαψοδία (which begins with a voiceless rho) . Thus, that is the interpretation we will adopt for Paonese. It is tempting to posit an analogous unvoiced ‘l’ based on the name “Coelho”- but as was noted in the overview, this is a straight borrowing from Portuguese, and so likely represents a fossilized spelling regardless of how the Paonese actually pronounce the name, which provides no good evidence.

We must still decide on the proper pronunciation of several other letters. As the English [ɹ] is relatively uncommon throughout the world, we will assume a trilled [r] for the <r> grapheme instead. The digraph <th> will be assumed always voiceless. The digraph <ph> will be taken as an arbitrary stylistic variation on <f>. Finally, for simplicity, vowels will take their IPA values, with the following exceptions:

1. The digraph “ae” will be taken to represent [e].
2. The letter <i> will be assumed to indicate a diphthong-forming /j/ whenever it follows another vowel.
3. The letter <y> will be taken an arbitrary stylistic variant for the diphthong /ai/.
4. The letter <c> will be taken as an arbitrary stylistic variant of <k>.
5. The letter <e> will be assumed to operate as an indicator of “long” values as is standard in English orthography when, according to my own intuition, it sounds good. In these cases, the <e> is silent.
6. other instances <e> will be taken either to represent /e/ (in open syllables) or /ɛ/ (in closed syllables¹³).

As previously noted, it is impossible to tell whether a particular unvoiced stop should be aspirated or not in certain environments; those cases are decided arbitrarily, so as to produce a reasonable mixture of aspirated and unaspirated examples. With issues of proper pronunciation decided, the complete corpus of attested Paonese words with transcriptions is as follows, with instances of arbitrary or exceptional decisions indicated in footnotes:

13 Closed syllables containing /e/ having been written with <ae> instead.

First Names

Beran	/ˈpe.ran/	Hessenden	/ˈhɛ.sɛn.tɛn/ ¹⁴
Can	/kʰan/	Mornune	/ˈmor.nu.ne/
Est	/ɛst/	Vilnis	/ˈvil.nis/
Gitan	/ˈki.tʰan/		

Last Names

Netsko	/ˈnet.sko/	Therobon	/ˈθe.ro.pon/
Paraio	/ˈpʰa.ra.i.o/		

Cities

Donaspara	/ˈto.na.spʰa.ra/ ¹⁵	Sherifte	/ˈʃe.rif.te/
Eiljanre	/ˈeɪ.lʒan.re/	Spyrianthe	/ˈspai.ri.an.θe/ ¹⁶
Koroi	/ˈkʰo.roi/		

Place Names

Cantatrino	/ˈkʰan.tʰa.tʰri.no/	Droghead	/trok/
Ferai	/ˈfe.raɪ/	Muniment	/ˈmu.ni.mɛntʰ/
Fraevarth	/ˈfre.varθ/	Pamalisthen	/ˈpʰa.ma.lis.θɛn/
Hyaline Gulf	/ˈhaɪ.a.lin/	Pergolai	/pɛr.'ko.lai/
Hylanth Littoral	/ˈhaɪ.lanθ/	Plarth	/plɑrθ/ ¹⁸
Hylanthus Sea	/haɪ.'lan.θus/ ¹⁷	Rovenone Canal	/ro.ve.non/
Jhelianse Sea	/ˈʒe.li.an.se/	Sgolaph	/zko.laf/
Lido	/ˈli.to/	Viamne	/ˈvi.am.ne/
Maesthgelai	/mesθ.'ke.lai/	Vredeltope	/ˈvre.tɛl.top/
Mathiole	/ˈma.θi.ol/	Qurai Peninsula	/qu.'raɪ/
Mervan Pond	/ˈmɛr.van/	Zelambre Bay	/ze.lampr/
Moravi Inn	/ˈmo.ra.vi/		

14 “Hessenden” represents a particularly special case; the initial <e> is not in a closed syllable when spoken, but given the conventions of English orthography, /he.sɛn.tɛn/ simply sounds indefensibly wrong. Note also that I have chosen to interpret the doubled <ss> strictly as an indication that the first <e> should not be interpreted as ‘long’ in the context of the second <e>. Cf. “Aiello”, where the doubled <ll> was interpreted as indicating a genuine doubled consonant, in the absence of any other obvious phonetic significance to that choice of English orthography.

15 Here, the ‘p’ actually lands in a logical location for secondary stress- one trochaic foot after the primary stress. Still, given the post-<s> positioning, adding aspiration is still somewhat arbitrarily.

16 Aspiration omitted arbitrarily.

17 /-us/ arbitrarily designated as a stress-attracting suffix, because that is how I pronounced it on my first time through the book.

18 Aspiration omitted arbitrarily.

Numerals

Ai	/aɪ/	Nona	/'no.na/
Shrai	/ʃraɪ/	Drona	/'tro.na/
Vida	/'vi.ta/	Hivan	/'hi.van/
Mina	/'mi.na/	Imple	/impl/

Continents

Aimand	/'aɪ.mant/	Nonamand	/'no.na.mant/
Shraimand	/'ʃraɪ.mant/	Dronamand	/'tro.na.mant/
Vidamand	/'vi.ta.mant/	Hivand	/'hi.vant/
Minamand	/'mi.na.mant/	Impland	/'imp.lant/

Miscellaneous

Auriol	/'a.u.ri.ol/	Mamarone	/'ma.ma.ron/
Ayudor	/'aɪ.u.tor/	mous	/'mo.us/
bogal	/'po.kal/	nli	/nli/
es	/ɛs/	Pao	/'pa.o/
en	/ɛn/	praesens	/'pre.sɛnz/ ¹⁹
furze	/furz/	Rhomel	/'rɔ.mɛl/
Kanetsides	/'kʰa.nɛt.saɪtz/	ro	/ro/

Now that we have the whole corpus in phonemic transcription, we can start doing some analysis to discover higher-level phonological rules and characteristics. The complete attested consonant inventory is as follows, with total number of occurrences listed beneath each phoneme²⁰:

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Velar	Uvular	Glottal
Plosive	p p ^h 12 3	t t ^h 24 4				k k ^h 7 4	q 1	
Nasal	m 24	n 48						
Trill		r r̥ 32 1						
Fricative		f v 5 11	θ 9	s z 14 5	ʃ ʒ 3 2			h 6
Lateral		l 21						

¹⁹ Aspiration omitted arbitrarily

²⁰ Note that, while transcriptions were of necessity done by hand, all of the statistics were automatically compiled by a short Python script, which was a major labor-saver. Learning a little basic computer programming comes highly recommended for the modern linguist, or conlanger.

There are a few gaps (e.g., no voiced /ð/, no aspirated /q^h/, no devoiced /l/ to match the /r̥/), but nothing glaringly weird, so I feel pretty good about declaring that this is the complete consonant inventory of Paonese. And while 68 words may not really be a statistically significant sample, we can get a general feel for appropriate relative frequencies of different phonemes- e.g., unaspirated stops are way more common than aspirated ones, /n/ occurs about twice as often as /m/, /q/ and /r̥/ each occur *about* once per 68 words (maybe a little more, maybe a little less)

For the vowel inventory, we have six pure vowels and three diphthongs:

a 48	ε 14	e 19	i 24	o 29	u 8
a ^j 15		e ^j 1		o ^j 1	

Once again, there's nothing particularly weird here, so I feel pretty good about declaring this the complete Paonese vowel inventory, with close-to-accurate relative frequencies.

Once we have some basic phonotactic / syllable structure rules, we could plug those frequencies into an automatic random word generator and start building a neo-Paonese dictionary. One approach to that which would vastly simplify the work of analysis would be to simply build a statistical model of frequencies for all attested pairs of phonemes, and all attested triples of phonemes, and then just feed that all into the random word generator; but on such a small corpus, that approach would be subject to severe overfitting. And besides that, it just feels like cheating.

To help figure out the syllable structure rules, I split all the words on (my arbitrarily decided) syllable boundaries, and then separately counted up all of the complete syllables, all of the unique onsets, and all of the unique codas. The frequencies for onsets and codas are as follows:

Onsets			Codas	
- : 21	k ^h : 4	pr : 1	- : 98	mp : 1
m : 20	s : 4	q : 1	n : 18	mpl : 1
n : 18	tr : 3	r̥ : 1	nt : 8	mpr : 1
r : 15	f : 2	sk : 1	l : 7	nt ^h : 1
l : 11	p ^h : 2	sp : 1	s : 5	nθ : 1
v : 10	ʃr : 2	sp ^h : 1	r : 4	nz : 1
t : 9	t ^h : 2	ʃ : 1	f : 2	p : 1
h : 6	ʒ : 2	t ^h r : 1	rθ : 2	rz : 1
p : 5	fr : 1	vr : 1	t : 2	st : 1
θ : 5	nl : 1	z : 1	k : 1	sθ : 1
k : 4	pl : 1	zk : 1	m : 1	tz : 1

A few things are immediately clear: Paonese *vastly* prefers open syllables, with no codas. While different from English, in which only about 40% of syllables are open, this is fairly normal cross-linguistically, and similar to languages like Italian and Spanish (Kubozono 1995). This may reflect some additional subconscious Italian influence on Vance's word-creation. More extreme cases are not uncommon- nearly 90% of all Japanese syllables are open, for example, and numerous languages allow only CV-structured syllables, of which 100% are open by definition.

Zero-onsets, in contrast, while certainly not rare in the Paonese corpus, are not that much more common than some of the other single-consonant onsets. There is also a lot more variety in possible onsets than there is in codas. Several phonemes (/ʒ/, /ʃ/, /h/, /q/ and /r/) can *only* occur in onsets. Also, there only two attested three-consonant clusters, and only in codas.

While the existence of coda consonant clusters technically place Paonese in the “complex syllable structure” category according to the World Atlas of Language Structures, it is definitely on the lower end, placing it close to the middle of the range of cross-linguistic variation (Maddieson 2013). With this level of syllable structure complexity, we would expect to see between 22 and 26 phonemic consonants in the phonetic inventory; the 20 phonemic consonants identified above would more typically be associated with a simple CV syllable structure, but this is still well within the range of known naturalistic variation. We can thus be fairly confident that my intuitive choices for syllable boundary placement were reasonable.

Unlike the basic phonetic inventory, I am not willing to accept that this data represents the complete set of all possible onsets and codas in Paonese, simply because the space of possible two-letter combinations is much larger than the space of single segments, and this corpus is far too small to be considered a representative sample of everything that Paonese might allow. We can use these frequencies for certain kinds of onsets and codas to improve a-priori word generation over what we would get from just sticking individual segments together with the right total frequencies; but, given that we have already accepted a certain minimum complexity in allowed Paonese syllable structure, some of these patterns should first be generalized.

Ignoring frequencies for now and looking only possibilities, reasonable generalizations for onsets seem to be that an onset can be formed from

1. any individual consonant.
2. any other consonant followed by an /r/ or /l/.
3. an /s/ or /z/ followed by any stop.

Codas, in turn, can be formed from

1. any individual coda consonant,
2. any coda consonant followed by /θ/,
3. any non-fricative coda consonant followed by /z/,
4. an /s/ followed by any non-fricative coda consonant,
5. a nasal followed by a homorganic stop (/t/, /p/, /t^h/ or /p^h/; no /k/, because there is no matching velar nasal),
6. any stop or homorganic nasal plus stop, followed by an /r/ or /l/ in word-final position,

where a coda consonant is any consonant other than /ʒ/, /ʃ/, /h/, /q/ or /r/. It may seem like the larger number of rules for codas vs. onsets contradicts the original observation that onsets have more variety, but note that the rules for codas are actually much more restrictive, allowing only specific special cases. Other rules that fit the data are possible, but these ones seem reasonable and aesthetically pleasing to me.

If we then examine the set of complete syllables, it turns out that there are four vowels that are used as complete syllables by themselves, with no onset or coda: /ai/ (3), /a/ (2), /o/ (2), and /u/ (2). One of these (/ai/) is attested as a complete word, while the other three at least occur in hiatus with other vowels on either side. The simplest generalization of this evidence is that any

vowel can form a complete syllable by itself.

It is theoretically possible at this point to look more extensively at the set of complete attested syllables, and try to come up with additional rules about what onsets and codas may be allowed with each other or with certain vowels, or what kinds of codas and onsets are allowed in hiatus within a word. But, the data is just too sparse to make really strong conclusions, so we might as well just say “anything goes”; once you have well-formed onsets and codas with vowels to put in between, they can be combined however you like. There are only three inter-syllabic restrictions which have been noted earlier, and which are restated and clarified here:

1. An /i/ cannot immediately follow any other vowel in the same word (with /a/, /e/, and /o/, it forms diphthongs, and combinations with other vowels just won't happen).
2. An onset /l/ is deleted if it comes in contact with a coda /r/.
3. Nasal+Stop+{/l/, /r/} codas can only occur word finally. If brought into contact with another syllable due to suffixing or compounding, they are re-syllabified, moving the /r/ or /l/ to the onset of the next existing syllable if possible and inserting an epenthetic /a/ to take the {/r/, /l/} onset otherwise.

The existence of /impl/ ~ /imp.lant/ suggests that maybe there should be a following-onset-deletion rule instead of the epenthetic /a/; however, I assume that we can treat “Impland” as an ancient fossilized form that does not necessarily reflect synchronic morphophonological principles, and deleting onsets feels like it would produce uncomfortable levels of homophony. The epenthetic vowel is chosen to be /a/ because /a/ is the most frequent vowel phoneme in the corpus by nearly a factor of two, and so seems like a good choice for a “neutral” vowel.

Sentences & Morphosyntax

Although, as previously mentioned, there is only one complete sentence of Paonese, there are two sentences glossed: the first actually provides the Paonese words, while the second has only the free English translation and the morpheme-by-morpheme gloss. The complete sentence is as follows:

Rhomel-en-shrai bogal-Mercantil-nli-en mous-es-nli-ro.

Statement-of-importance - in a state of readiness - two; ear-of Mercantil - in a state of readiness; mouth - of this person here – in a state of volition

"There are two matters I wish to discuss with you."

It is claimed that "The italicized words represent suffixes of condition." I suspect italicizing *Statement-of-importance* is simply a typographical error in the text of the book, as it is clearly not a suffix, being the first thing in the utterance.

This first gloss gives us some insight into morphological structure and some interesting morphemes. The word *shrai* for “two” we already know. *Rhomel* could be interpreted as a grammatical morpheme for emphasizing the importance of an utterance, but it seems to make more sense as a noun for “something important”. The word *bogal* clearly means “ear”, while *mous* must be mouth. *Mercantil* is a proper name referring to the addressee of this utterance. We also have the following grammatical morphemes:

<i>-en</i>	"readiness"	<i>-nli</i>	genitive
<i>-ro</i>	"volition" (agent)	<i>es</i>	"this person"

The "readiness" marker might be an existential, or perhaps a topic marker, or it could literally mean only "ready for something". The morphemes *-en* and *-ro*, described as *suffixes of condition*, are the best matches we have so far for the expected "post-positions"; they seem to have functional, adpositional meanings, and the fact that every noun seems to require one matches the description of postpositions as main component of the language. However, given that they are written attached to the previous root, and can have additional material appended to the same orthographic word, it makes most sense to abandon the postpositional analysis for now and instead simply call these suffixes, as explained along with the gloss as it appears in the text.

The only direct evidence we have for what kind of material can be appended to a noun-condition complex is a simple numeral, in this case the number "two"; but, it seems reasonable to assume that the closed class of simple numerals can probably be extended with morphemes for, e.g., "few", "none", and "many" to allow for the expression of non-specific plurality as desired, without having number as an obligatory grammatical category. There's no indication of how larger numbers are formed, however, and those often have different grammatical behavior in natural languages.

Genitives apparently follow their heads in compounds and use a postposed genitive marker *nli*.

It is particularly interesting that *es*, the obvious translational equivalent for "I", is glossed as "this person here" (since we know in context that Mercantil is the addressee, we can assume that that is the source of the pronoun "you" in the English translation). There isn't any explicit discussion of Paonese pronouns²¹, so we cannot tell whether there is some other word that means only "I", or any other specific personal pronouns. Given what we know of the highly communal, non-individualistic nature of Paonese culture and the extra-fictional knowledge that the original Paonese as well as the artificial languages that replace it are meant to reflect the strong Sapir-Whorf hypothesis, it is reasonable to assume that Paonese in fact does not have any explicit personal pronouns. In that case, *es* and other translational equivalents for English pronouns may in fact have more in common with Japanese pronominals; i.e., while conventionally used for "I", *es* does not literally mean *only* "first person singular", and can be employed as a more general noun, most likely referring to any nearby or highly relevant person in the proper discourse context. We know, however, that *es* is almost certainly exclusively animate, if not exclusively human, and thus we can predict an animate/inanimate split in the set of conventional pronominal words.

The second gloss we have is as follows:

"The farmer chops down a tree."

Farmer *in a state of exertion*; axe *agency*; tree *in a state of subjection to attack*

Contrasting with the previous gloss, we can determine that, although we do not know its form, there is some suffix for physical exertion as distinct from simple will, which may justify the development of Paonese specific terminology of "volition" for the first case as opposed to a

21 There is, however, a short discussion of Breakness pronouns, Breakness being the fictional natural language of the eponymous planet in Pao's universe. Said discussion serves mainly to explain that the people of Breakness are so egocentric that there's no need to bother with an explicit word for "I".

more generic “agent”. *Agency* in this case, however, almost certainly refers to an instrument in more standard terminology, and is evidence of another “suffix of condition” for which we do not know the actual form. Further, there is a third unknown condition suffix for being subject to attack. This seems strangely specific; it is possible that this is a misleading gloss by the naïve narrator, but things get more interesting if we assume in this case that the unknown suffix really means only and specifically that. That level of specificity suggests that “suffixes of condition” are something like the lexical suffixes in native languages of the American Pacific Northwest. Unlike lexical suffixes in Native American languages, however, these refer not to concrete objects, but to verbal concepts. It is possible that the “suffixes of condition” may not even be a closed class of grammatical morphemes. In contrast, they could be an open class of always-intransitive verb-like morphemes which are merely required to occur bound in a compound with a noun.

In any case, these are definitely not complete verbs as they are usually understood; no single “suffix of condition” controls the meaning of the entire sentence. Rather, the “picture of a situation” that Paonese is said to represent is built up by contributions from all of the “suffixes of condition” on all of the nouns in a clause. Leaving the question of whether they form an open or closed class undecided for the moment, we will from here on adopt the term “semiverbs” to describe these components of Paonese morphosyntax.

As an incidental cultural note, we also now know that Paonese has *some* word for each of “farmer”, “tree”, and “ax”.

Tense & Aspect

In addition to nouns and postpositions, we are told that Paonese contains “temporal indices” as a major component. Unfortunately, the available example sentences examples leave us with no direct evidence as to what “temporal indices” might actually be; there’s nothing that might match that description in either of the given glosses, which indicates that whatever they are, they must be optional rather than obligatory components of a grammatical sentence. It is unlikely that they were simply omitted from the gloss as unimportant or hard to translate. According to the narrator, the sole point of providing the Paonese words for gloss #1 was to demonstrate that it isn’t substantially longer than the equivalent English utterance, a purpose which would not be properly served by lying about which words should actually be there. The known glosses, and the general static nature of Paonese culture in light of the strong Sapir-Whorf hypothesis, suggest that Paonese probably does not have grammatical tense. That does not, however, rule out optional morphemes with temporal reference that could pass as “temporal indices”. Plenty of natural languages get by just fine with no grammatical tense marking, while still being able to express time when needed by the use of explicit adverbs (like “yesterday”, “tomorrow”, or “right now”) or other temporal constructions. Mandarin Chinese, for example, works this way.

Mandarin Chinese does, however, have grammaticalized aspect. Where *tense* tells you *where* an event occurs in time, *aspect* relates to *how* an event is embedded in time. The two concepts are frequently conflated in natural languages, both in terms of how they are marked and how speakers conceive of their verbal systems, so it is possible that “temporal indices” also have something to do with aspect.

Furthermore, despite the fact that Jack Vance was untrained as a linguist, the concept of “tense” should nevertheless be familiar to any educated layman. If, therefore, he meant

“temporal indices” to refer to tense markers, we thus have to wonder why he did not just *say* “tense”; this line of reasoning also points towards an aspectual interpretation. One fairly straightforward explanation may be that he simply recognized that “tense” is not the name of a part of speech as the word is typically used in English grammar; and rather than confusing the reader by re-purposing it as such, he invented a novel term. But even in that case, we have to realize that “tense” and “aspect”, while formally separate grammatical categories, are extensively conflated in many languages, and extensively confused by linguistically-naïve English speakers.

As a slight digression, we might briefly consider what Paonese could be like if “temporal indices” do strictly encode tense. If Paonese still had aspectual distinctions, but not in the “temporal indices”, an alternative possibility would be that Paonese have *lexical aspect*, like many Slavic languages. In a language with lexicalized aspect, verbs come in sets of multiple words, often morphologically derived from each other but not necessarily, which all have the same basic meaning but encode different aspects. This would be consistent with the lack of explicit aspectual morphemes in the two attested Paonese glosses, but seems difficult to arrange when the description of an event is spread out over multiple semiverbs- which one would control the aspectual marking? Due to the lack of coherent verbs to carry aspectual marking in Paonese, it seems that any aspectual marking would *have* to be carried on some other distinct kind of word, which provides further evidence for an aspectual interpretation of “temporal indices”.

Finally, while at first glance the term “index” seems to me to fit more closely with the idea of words that specify specific places in time (i.e., tenses), it also makes sense that these “indexes” may be meant to indicate only relative ordering of events within a single discourse, and not necessarily absolute position with respect to the time of utterance; and that interpretation actually fits better into the Whorfian mold, in which the Paonese language reflects the Paonese culture, unconcerned with its placement in time. This kind of “relative tense” can be encoded in a language’s aspectual system, orthogonal to tense, as indeed it is in English, via the perfect aspect constructions with the auxiliary verb “have” (e.g., “to have gone”- a tenseless perfect-aspect infinitive, as compared to “go” or “went”- the finite present and past tense forms). Due to this fact and our knowledge of the ease with which an anglophone narrator could conflate the concept of aspect with tense, we will therefore settle on the idea that “temporal indices” are optional adverbial words which primarily encode aspectual information (though some indication of non-grammaticalized tense cannot be completely ruled out).

Replacing Adjectives

The fact that Paonese is claimed to have no adjectives is comparatively unremarkable and unproblematic. Plenty of natural languages get by with no distinct class of adjectives, making use of nouns or verbs to fill that role. Paonese has no easily-identifiable verbs, so nouns it will be. We already have direct evidence that multiple noun roots can be compounded together in the presence of the genitive marker *nli*. Keeping with the intended polysynthetic nature of the language, it is a small step from there to assume that there may be other combining markers to indicate various descriptive relations between different roots in a larger compound. We can further allow for arbitrary sequences of unmarked noun roots in which each component is interpreted as an alternate descriptor of the same referent, taking the intersection of all of their individual meanings.

The combining markers like *nli* have many of the qualities of adpositions, as semiverbs initially seemed to. If we were to insert word boundaries after the combining markers around noun roots (an arbitrary decision as Paonese has no morphophonological evidence to definitely

distinguish two words in hiatus from two bound components of a single compound), they would appear exactly like postpositions, with the exception that postpositional phrases seem to follow rather than precede their heads; this is typologically *unusual*, especially when the evidence of postpositions and semiverbs together indicates the Paonese is generally head-final, but it's hardly impossible.

Combining markers are quite distinct from semiverbs in that they specify the relation of nouns to other nouns, not the relation of a noun to a predicate (or “situation” in Paonese terms). However, since English does not distinguish adpositions that modify nouns from adpositions that modify verbs or clauses, and some languages have precisely the opposite restriction (i.e., that adpositional phrases only or at least primarily modify clauses, not nouns)²², it is not at all surprising that a naïve narrator could confuse the two, leading to the claim that “post-positions” are a primary component of the language while ignoring the distinction between these combining morphemes and the equally essential semiverbs.

This style of compounding allows for the translation of fairly complex noun phrases, including arbitrarily long sequences of adpositional modifiers and arbitrarily long strings of adjectives at every level. There is some room for structural ambiguity in terms of whether a second adpositional modifier attached to a first adpositional modifier, or to the head of the whole phrase, but that's no different than what we already have to deal with in English²³, and there's really no reason to ask for any more in order to get a fully functional, fully expressive human language. While this was my final decision regarding Paonese, however, I must note that more precise systems, with no such structural ambiguity, are possible at very little extra cost in complexity. In particular, I have been inspired by a four-category compounding system described by Herman Miller in a message to the CONLANG-L mailing list²⁴, implementable in either morphology (for building polysynthetic compound words) or syntax (for building complex noun phrases), which allows for arbitrarily complex trees of heads and modifiers. To a native English speaker like me, it gets a little confuddling fairly quickly, but one can imagine that perhaps a native speaker of Inuktitut or Turkish would not find these arbitrarily complex compounds quite so daunting. Expanding the combinatorial system of Paonese to accommodate such things would be a perfectly valid alternate choice for a different reconstruction.

Adverbs

Having addressed the issue of adjectives, the logical next step is dealing with adverbs. Official descriptions of the language do not mention adverbs, but do not rule them out either. If we are correct about the nature of “temporal indices”, those roots would be effectively adverbial in nature. Although we have no direct evidence of other adverb roots, it is easy to argue that Paonese sentences so far are built from nothing *but* adverbs; each noun+semiverb complex is effectively an adverb which adds a little bit more to the description of a situation.

Given that starting point, we should expect the evolution over time of unanalyzable adverb roots from what were originally idiomatic noun+semiverb complexes, filling out the class

22 Standard practice in Classical Latin pedagogy, for example, is to claim that adnominal usages of prepositions are strictly disallowed (Wharton 2009, 184). While the true situation is somewhat more complex, adnominal prepositions are less common in Classical Latin than are adverbials.

23 E.g., in “the vine in the pot on the mantle”, is the pot on the mantle, or is the vine merely draped over the mantle, with no clear indication of the specific location of the pot? If that seems like a distinction without a difference to you, something that would hardly ever actually matter- well, you're probably right. Which is why we won't worry about it in Paonese, either.

24 <https://listserv.brown.edu/archives/cgi-bin/wa?A2=CONLANG;9a6d18ef.1306D>

of distinct adverbs. “Temporal indices” may have been singled out for special mention, in preference to a larger class of adverbs, due to a combination of factors, including the relative difficulty of distinguishing many other adverbs from simple noun+semiverb complexes and/or a relatively high frequency of usage compared to other semantic classes of adverbs.

Given that adjectival concepts are represented in Paonese by noun compounding, certain adverbial concepts that would normally show up as adjective modifiers in English²⁵ could potentially also be handled by additional derivational morphology on noun roots.

The Unknown & The Unknowable

We are now at the very edge of what we can reasonably tie to evidence from the book. There are many topics still to be covered and decisions to be made to fill out a complete description of a language, but they will need to be made on a fairly arbitrary basis.

For example, how do we fill out the pronominal system? It was previously noted that Paonese may behave much like Japanese in not having a class of strict pronouns distinct from normal nouns; but even so, we have to ask: what kinds of words do they use for those functions? Just running with the gloss “this person” for *es*, I see opportunities for three different three-way distinctions: three persons, three animacies (human, non-human, inanimate), and three degrees of proximity (this, that, yonder). We can assume that non-human and inanimate first person don’t make much sense; nor does inanimate second person. That takes three cells out of the chart, but still leaves 24 possible personal and demonstrative pronominal forms, and room for plenty of other irregular additions in the Japanese style. Is that the *best* way to reconstruct Paonese pronouns? I don’t know, but it is not contradicted by the evidence.

We also have no evidence for how complex sentences are structured. Does Paonese have nominalized clauses? Relative clauses? Complementizers? Coordinating or subordinating conjunctions of any kind? All of this will need to be invented from scratch.

There does not seem to be any strong reason to assume that Paonese does not have coordinating or subordinating conjunctions. They may not be in the list of “nouns, post-positions, and temporal indices”, but if you had to describe the major components of English, would you include conjunctions, or prefer a list like “nouns, verbs, and adjectives”? So we might as well posit a fairly normal array of useful conjunctions. Given the lack of verb phrases or other clear clausal structure, however, it feels right to say that Paonese should have separate classes of conjunctions that join noun+semiverb complexes and conjunctions that join separate clauses, thus helping you identify where one “situation” ends and another begins. Armenian is an example of a natural language with just such a distinction.

The decentralized nature of the Paonese clause and the way nominal arguments are incorporated into verb complexes also make me feel like the clause structure should be relatively flat- i.e., there should not be embedded relative clauses or nominalized complement clauses. How, then, would those concepts be handled by Paonese, if we needed to translate them? Well, instead of “I saw the man who had a telescope” (for example), you would just have to say “I saw the man and he had a telescope.” That doesn’t sound like a very felicitous paraphrase in English, because the pragmatic implications are all wrong- but there’s no reason that should have to be the case in Paonese. For complement clauses, we can simply state the complement and then refer to it as “that” later on; e.g., “I know you like ice cream” could be “You like ice cream, I know that.” This is a case where an explicit complementizer could be useful to mark the boundary between

25 E.g., “very”, although that particular example may be marginal due to the need to avoid overt comparison.

clauses, but intonation or a regular conjunction would serve just as well.

All in all, I don't think Paonese needs complementizers. If it had them, they would probably come at the end of a clause, given the general head-final nature that we have already observed. Perhaps some of the inter-clausal conjunctions derive from intra-clausal conjunctions, postpositions, or semiverbs that fused with the complementizer of the previous clause in an earlier form of the language; that is something that might be taken into account when generating additional vocabulary. But, in the synchronic analysis, clause structure can easily be completely flat, and this seems to better fit the intended *feel* of the language- not only does it suppress explicit comparisons, even the syntax banishes any hierarchical ordering of clauses and any idea of "subordination" to a superior.

And that brings us to the last unknown: what are all the rest of the Paonese words? How do we fill out the vocabulary? If we were reconstructing a lost natlang, we would try to get samples of its relatives, and do comparative analysis to figure out what the equivalent words in this language should be. But even if we had knowledge of related languages, the number of words that are *not* proper names or grammatical morphemes, for which we could try to find reasonable correspondences, is vanishingly small- just the three of "rhomel", "bogal", and "mous", and perhaps "Ayudor". Even including the rest of the corpus, there's barely enough there to plausibly isolate a couple of re-used morphemes. And we do not, in fact, have any real knowledge of any related languages. The use of French, Italian, and Portuguese names suggests that perhaps we should consider the use of Romance roots. But, what we do know is that Paonese is not a direct descendent of any current Romance language; its evolution was, at the very least, filtered through the intermediate language Waydalic, in ways that are completely unknowable; and what little we do know suggests that any original Romance influence would be completely unrecognizable in the language now. More likely, these Romance elements were introduced as a superstratum influence on an already-existing Paonese, not as ancestral holdovers. But that, at least, is something- we can plan to borrow large swaths of Romance vocabulary into Paonese grammar for science, technology, and "high class" registers, which greatly reduces the work that must be put into strictly a-priori vocabulary generation. Even so, if anyone else does an independent reconstruction of a different potential Paonese, this is where it is likely to diverge the most- even if you come up with the exact same phonetic inventory, and the exact same set of phonotactic rules, for all of the thousands of common words that you would need to add to make Paonese a complete usable language, you'll just have to make them up yourself.

Acknowledgments

First of all, a reconstruction of Paonese, and an article about it, could not possibly exist without the work of author Jack Vance to provide the raw materials. Mr. Vance died in May of 2013, while I first started working on my reconstruction in July of that same year- only a few months too late to be able to ask questions of the only authoritative source! While my timing was, in honesty, coincidental, some portion of this work is therefore dedicated to his memory.

I am also indebted to the reviewers who helped me figure out how to turn my jumble of nonlinear thoughts into a reasonably well-organized document, and provided several additional insights along the way which made it back into my reconstruction.

And finally, thanks to the CONLANG-L mailing list and all of its members in general who have served as a great support to my love of conlangs and education in linguistics over the years, as well as serving as a great sounding board for ideas during this and other projects.

References

- Kubozono, Haruo. 1995. "Perceptual evidence for the mora in Japanese". In *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*, 141-157. Cambridge, UK: Cambridge University Press
- Le Guin, Ursula K. 1974. *The Dispossessed*. New York: Harper Collins.
- Maddieson, Ian. 2013. "Syllable Structure". In *The World Atlas of Language Structures Online*, Dryer, Matthew S. & Haspelmath, Martin (eds.) Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed Oct. 2015. <http://wals.info/chapter/12>
- Wharton, David, "On the Distribution of Adnominal Prepositional Phrases in Latin Prose", *Classical Philology* 104 (2009): 184-207, doi: 10.1086/605342
- Williams, David B. 2006. "[Vance Museum - Miscellany - Biographical Sketch](http://www.vancemuseum.com/vance_bio_1.htm)". Accessed Sept. 2015. http://www.vancemuseum.com/vance_bio_1.htm
- Vance, Jack. 1958. *The Languages of Pao*. New York: Avalon Books
- Yavaş, Mehmet. 2012. "English Consonants". In *Applied English Phonology, Second Edition*, 57-73. Chichester, UK: Blackwell Publishing