Universal Dependencies in a galaxy far, far away...

What makes Yoda's English truly alien

Natalia Levshina
Leipzig University
IPF 141199 Nikolaistraße 8-10
04109 Leipzig
natalia.levshina@uni-leipzig.de

Abstract

This paper investigates the word order used by Yoda, a character from the Star Wars universe. His clauses typically contain an Object, Oblique and/or non-finite part of the predicate followed by the subject and the finite predicate/auxiliary/copula, e.g. *Help you it will*. Using the sentences in Yodish from the scripts of the Star War films, this paper examines three cross-linguistically common tendencies, which can be explained by optimization of processing: the trade-off between entropy of S and O order and morphological cues, minimization of dependency lengths, and the tendency to place the verb in the end of a clause. For comparison, a standardized version of Yoda's sentences is used, as well as the Universal Dependencies corpora. The results of quantitative analyses indicate that Yodish is less adjusted to human processor's needs than standard English and other human languages.

1 Introduction

It is well-known that some word order patterns are more likely to occur than others. For example, the overwhelming majority of the world's languages tend to place subject before object (Tomlin 1986; Dryer 2013). There are also more subtle tendencies, which become visible only with the help of quantitative analyses. For instance, it has been shown that language users tend to minimize the average syntactic dependency lengths (Futrell et al. 2015). These and other tendencies have been explained by optimization of processing and human cognitive biases. In particular, the dependency length minimization has been explained by the limitations of human working memory (Gibson 1998; Hawkins 2004; Futrell et al. 2015), whereas the preference for placing subject before object has to do with the semantic and pragmatic properties of the main arguments, and the fact that topical and animate arguments, which are usually subjects, are more likely to appear first (Tomlin 1986).

These explanations are based on specific assumptions about human cognition. But what if the language user is an alien? The present study investigates the properties of word order observed in the speech of Yoda, a powerful Master Jedi from the Star Wars universe. Yoda appeared in most of the films of the franchise (Episodes I, II, III, V and VI, as well as the sequels *The Force Awakens* and *The Last Jedi*, as a voice). Yoda belongs to an unknown species. One of his distinctive characteristics, in addition to large green ears, is the use of unusual word order patterns. Some examples are provided below:

(1) Friends you have there. (E V)[1]
(2) Help you it will. (E II)
(3) The secret of the Ancient Order of the Whills, he studied. (E III)

[1] 'E' stands for Episode.

Yoda's version of English will be called Yodish in this study. But why should a study of an artificial language like Yodish be interesting to syntacticians? These data allow us to understand better the properties of human languages, in particular, what is possible and impossible, likely and unlikely, efficient and inefficient. In particular, we will focus on the following manifestations of word order efficiency:
1) the implicational relationship between lack of rich morphology (especially case system and agreement) and rigid word order (e.g. Sapir 1921; Kiparsky 1997);
2) the universal tendency towards minimization of dependency lengths (Hawkins 2004; Futrell et al. 2015);
3) the preference for the final position of the verb in a clause. This tendency has been explained by the efficient strategy of putting verbs, which require more cognitive efforts in terms of categorization and acquisition than nouns, in the position where the former are more predictable and therefore easier to process (Ferrer-i-Cancho 2017).

In order to investigate whether these cross-linguistic tendencies are also observed in Yodish, we use the film scripts of five Star Wars films parsed according to the Universal Dependencies annotation schema. These data are compared with the Universal Dependencies corpora (Nivre et al. 2017) and with a standardized version of Yodish, created manually by myself.

The remaining part of the paper is organized as follows. Section 2 presents the data and describe the properties of Yoda's word order in Section 2. Next, Yodish is compared with human languages, using the Universal Dependencies corpora and the normalized version of Yoda's sentences in standard English. In particular, the implication relationship between rich morphology and rigid word order is discussed in Section 3, followed by a study related to minimization of dependency lengths (Section 4), and the position of verb in a clause (Section 5). Section 6 provides a summary and a discussion of the findings.

2 Yodish: data and main word patterns

The Yodish data have been collected from the Internet Movie Scripts Database.[2] Data from five episodes were used: two episodes from the original trilogy (Episodes V and VI) and three episodes from the prequel trilogy (Episodes I, II and III). One-word utterances and paratactic structures were excluded as irrelevant for the study. In total, the sample contained approximately 2000 words.

Yoda's word order has been described as some linguists as OSV or XSV, where X stands for any complement that goes with the verb.[3] However, it seems that a more precise description would be as follows:

(4) Non-finite part of predicate/Object/Oblique Subject Finite Verb/Auxiliary/Copula

The first part can be object, oblique, nominal part of the predicate or non-finite parts of the predicate, i.e. participle or infinitive with dependent elements. They are followed by the subject and the finite verb, auxiliary or copula. Below are some examples that support this interpretation:

(5) Rest I need (E VI)
(6) To his family, send him. (E III)
(7) A certainty it is. (E II)
(8) Hard to see, the dark side is. (E I)
(9) Earned it, I have (E VI)

If the non-finite lexical verb has arguments, they usually follow the verb, as in (10):

(10) **Save them**, we must. (E III)

---

[2] See https://www.imsdb.com/ (last access 29.04.2019).
[3] See a blog post by Geoffrey K. Pullum at http://itre.cis.upenn.edu/~myl/languagelog/archives/002173.html, a post by Mark Libermann at http://itre.cis.upenn.edu/~myl/languagelog/archives/002182.html, and an article in The Atlantic by Adrienne Lafrance https://www.theatlantic.com/entertainment/archive/2015/12/hmmmmm/420798/ (last access 29.04.2019).

But occasionally an argument may precede the verb, as in (11):

(11) Master Kenobi, **our spies contact**, you must, and then wait. (E III)

The subject is usually followed by the finite form, including copulas and auxiliaries, as in (12), although some sentences have the reverse order: first copula/auxiliary, and then subject, as in (13):

(12) Subject – copula/AUX:
Unexpected **this is**, and unfortunate. (E VI)
Earned it, **I have** (E VI)

(13) Copula/AUX – subject:
Not ready for the burden **were you**. (E VI)
Heard from no one, **have we**. (E III)

In addition, there are rare cases when the non-finite verb is in the end, after the modal or auxiliary:

(14) The outlying systems, you must **sweep**. (E III)

The most substantial source of variation, however, is the frequent use of conventional SVO word order, as in the examples below:

(15) Master Obi-Wan has lost a planet. (E II)
(16) A Jedi's strength flows from the Force. (E III)
(17) That place is strong with the dark side of the force. (E V)

Compare two very similar sentences with different word order:

(18) Reckless is he. (E V)
(19) You are reckless (E V)

It is not clear what may condition this variation. In the absence of evidence to the contrary, we assume that Yodish is a special variety and does not represent code-switching between standard English and Yoda's own dialect.

One should mention that OSV is not entirely alien to English speakers: it is permissible in certain emphatic contexts, as the one below:

(20) **This** I could **understand** but why not tell me this some three and a half hours earlier.[4]

However, the structures in Yodish are not emphatic or contrastive like (20).

Remarkably, the films vary substantially in terms of the deviations of Yoda's word order from standard English. In order to quantify the differences, the original (Yoda's) token IDs were compared with the IDs in the normalized version, re-written in standard English, and the proportions of discrepancies for each of the five films were computed. The original trilogy has lower number of words that deviate in their order from standard English: Episode V has only 25%, while Episode VI has 36.1%. The prequel films have higher numbers. In Episode I, the difference is 56.3%; Episode II has 51.2%. Finally, Episode III has 67.4%, the highest number of deviations.

---

[4] URL https://www.nhs.uk/Services/hospitals/ReviewsAndRatings/DefaultView.aspx?id=95110&SortType=1,5,5,5,5,5,5&pageno=3&subject=All%20subjects&spid=0 (last access 29.04.2019).

It is interesting that the events of Episode V happen after Yoda was in exile for many years. Therefore, one would expect that his solitude would lead to less standard English, due to the lack of language contact. However, the opposite is the case. This counterintuitive fact can be explained by Yoda's character development, which does not correspond to the timeline in the Star Wars universe. First, the original Episodes V and VI were created, and only later the prequel films with Episodes I to III. As Yoda's film character was developed by the authors (primarily, by George Lucas), the divergent word order became more and more frequent.

3 Trade-off between word order entropy and morphology

Recognition of the core arguments, or who did what to whom, is crucial for interpretation of a sentence. Rigid word order can compensate for the lack of morphological cues that help to recognize the main arguments, as in English or Mandarin Chinese. Flexible word order is observed overwhelmingly in languages with rich morphology, such as the Slavic or Baltic languages. This relationship, which is well-known in linguistics (Sapir 1921: 66; Jakobson 1936[1971]: 28; Blake 2001: 15), has been formulated as a one-way implication by Kiparsky: "lack of inflectional morphology implies fixed order of direct nominal arguments" (Kiparsky 1997: 461; see also McFadden 2003: 301). Similar claims have been supported by typological evidence (Siewierska 1998; Sinnemäki 2008), corpus data (Futrell et al. 2015) and experimental evidence (Fedzechkina et al. 2016).

Yodish, like English, has poor inflectional morphology and no case distinctions for nominal subject and object. Therefore, it must have a strong preference for a specific ordering of subject and object. In order to see if this prediction holds, Shannon's entropy was computed using the standard formula:

(21) $H(X) = -\sum_{i=1}^{2} P(x_i)\, log_2\, P(x_i)$

where X is a binary variable representing two possible word orders, Subject + Object (SO) and Object + Subject (OS). P (x_i) is the probability of one of the orders, which equals its relative frequency (proportion) in the corpus. If the proportion of SO is 1, and the proportion of the reverse order OS is 0, or the other way round, the entropy H is equal to zero. That is, there is no variation. If the proportion of each of the possible word orders is 0.5, the entropy takes the maximum value of 1. If both orders are attested, and one of them is more frequent than the other, then H lies between 0 and 1.

The frequencies of SO and OS were collected manually, since the number of sentences with an explicit nominal or pronominal subject and object is rather low. The results for each film are shown in Table 1. Only the main clauses were taken into account because the order of subordinators (complementizers and relativizers) is determined by other syntactic factors. Non-finite clauses and questions were excluded, as well. The entropy values are high. One can see that there is no strongly preferred order. Even though some of the frequencies are rather low, the overall pattern is quite consistent. The highest entropy is observed in Episode V, the first film where Yoda appears, and the lowest in Episode I (which may be due to the small number of examples).

Film	Frequency SO	Frequency OS	H
Episode I	2	7	0.76
Episode II	3	8	0.85
Episode III	7	18	0.86
Episode V	18	15	0.99
Episode VI	3	6	0.92
Total	33	54	0.96

Table 1: SO and OS frequencies in Yodish and their entropy.

Let us now compare these results with human languages. Figure 1 displays the entropy values of subject and object (co-dependents) in Yoda’s language and in the languages of the Universal Dependencies corpora (version 2.2). Only the main clauses are considered, since the object and subject in subordinate clauses can be expressed by subordinators, which influences their position. Yodish is represented by separate values for each film, plus the value Yoda_av computed on the total frequencies in all films. The Yodish values are displayed as black dots. In the UD corpora, only the corpora with the total number of subject – object pairs above 50 were considered.

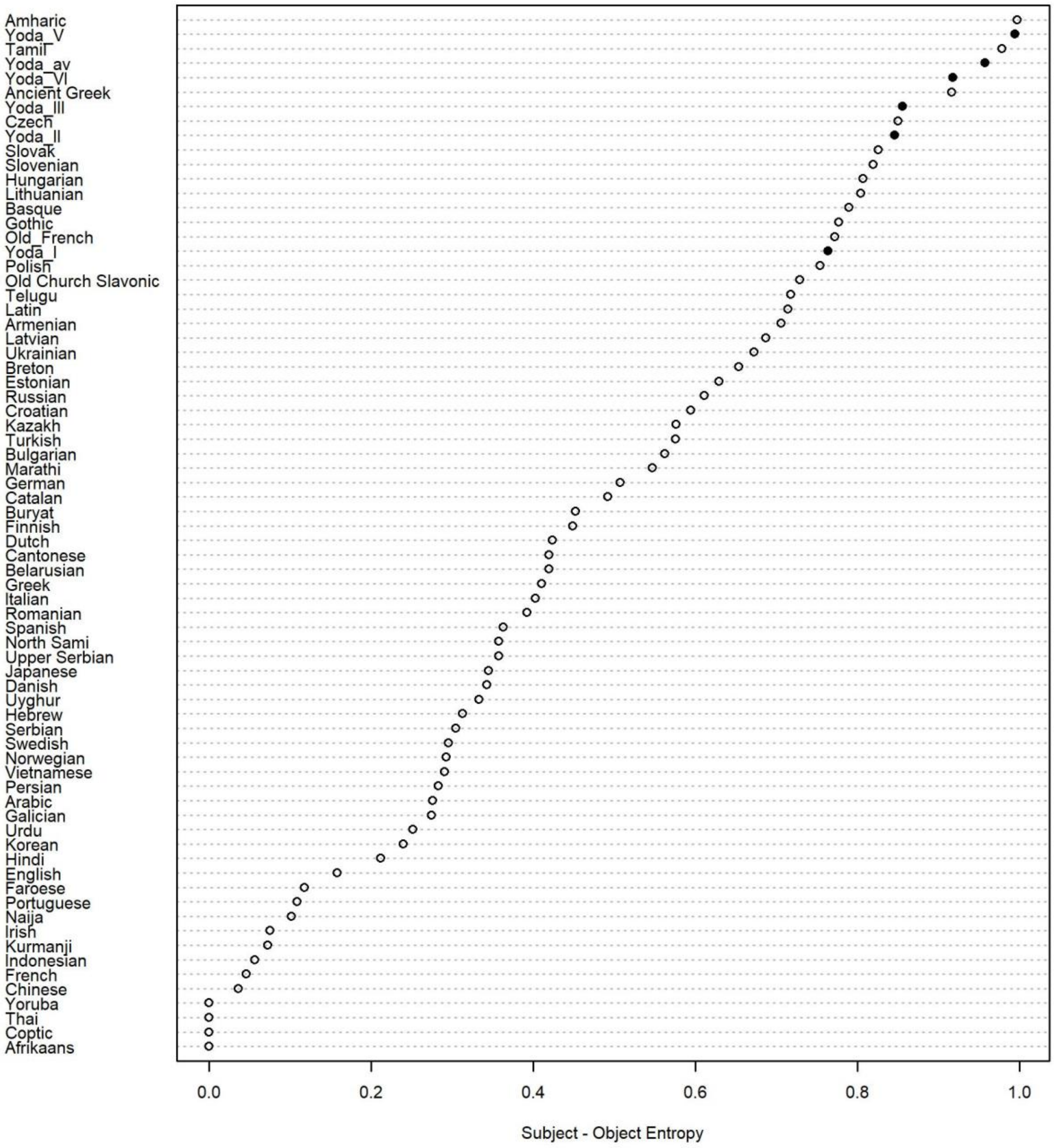

Figure 1: Subject-Object Entropy in Yodish and UD corpora.

According to the data, the highest entropy is observed in Amharic. However, this result seems to be an artefact of some annotation decisions made by the corpus creators. Amharic has subject clitics, which follow the verb to mark agreement. An inspection of a corpus sample reveals that the clitics are annotated in different ways depending on the presence of a full nominal subject. If there is no full nominal subject, they have the tag of subject ('nsubj'). In the presence of a nominal subject, they are annotated as expletive elements ('expl') – i.e. nominals that appear in an argument position of a predicate but which do not themselves satisfy any of the semantic roles of the predicate. This explains the high variability of the subject and object word order in Amharic.

The next lines are occupied by some variants of Yodish, with Tamil, Ancient Greek and some Slavic, Baltic and other synthetic languages in-between. Thus, Yodish behaves like the languages that have rich morphology and case marking. In contrast, English and other analytic languages are located at the bottom because they have low entropy values. From this follows that Yodish is truly alien. The lack of syntactic or morphological cues means that the hearer needs to invest extra processing efforts in order to understand who did what to whom.

4 Dependency lengths

This section addresses the question whether Yodish is efficient with regard to dependency lengths. As was already mentioned in Section 1, human languages minimize the distance between the heads and dependents. This has been explained in terms of processing efficiency: long dependencies are more difficult to process. When the head and dependent are located far away, it creates extra load for memory because of the long timespan over which the head or the dependent must be held in a memory store (Hawkins 2004; Futrell et al. 2015).

It does not seem very reasonable to compare dependency lengths across different languages because the number of words depends on the properties of a language (e.g. synthetic or analytic) and orthographic conventions. In order to have a basis for comparison, the Yodish data were normalized such that the order was like in standard English. Sentences that were difficult to transform into normal English were excluded. So were one-word replies and parataxic structures. In total, the dataset had 305 sentences in both orders.

Next, the sentences were parsed syntactically with the help of the R package udpipe (Wijffels 2018) according to the Universal Dependencies annotation style based on the UDPipe software (Strakov and Straková 2017). A manual check of each sentence was performed, and the parsing errors were corrected. After that, dependency lengths were computed for the normalized version and the original Yoda's version with the help of an R script. The punctuation marks were not taken into account in the computation of lengths, i.e. they were not regarded as full-fledged tokens with their own ID. Consider an example:

(22) a. **Heard** from no one, **have** we. (E III, original)
b. We **have heard** from no one. (standardized)

If we take the relationship between the auxiliary have and the participle heard, the distance in the original (22a) is 4 words (the comma is excluded). In the standardized version (22b) it is only 1 word.

It is necessary to mention here that there exist different approaches to deciding on what should be regarded as the head and what as the dependent. For example, many syntactic theories say that prepositional objects depend on their prepositions, whereas the UD corpora annotate prepositions as dependents of head nouns, marking the former as a case relation. Futrell et al. (2015) compared different approaches are concluded that the results were the same, as far as the tendency to minimize dependency lengths is concerned.

The dependency lengths of the terms of address, coordinated and parataxic sentence parts, as well as compounds (mostly nominals) were excluded from the subsequent aggregate analyses because the notion of head is difficult to apply there. The roots, which have no lexical head, were excluded, as well.

The quantitative analyses reveal that Yodish has on average longer mean dependencies than standard English. Consider the average lengths presented in Table 2. The smallest difference is in Episodes V and VI – this is due to their more standard character in general, as was already mentioned. It is interesting that the other word order is observed mostly in short sentences. In longer sentences, there is no difference. This makes the difference in dependence lengths less observable.

A series of the Wilcoxon's paired rank-sum tests (where two measures are provided for each word in the text) performed on each film demonstrate that the difference is only statistically significant in Episodes I and III, and marginally significant in Episode II. In the original trilogy, the difference is not statistically significant.

Note that Episode VI, in which Yoda dies, only has 29 analyzable sentences, so the test may not have sufficient power due to the small sample size.

Film	number of comparable sentences	mean length in original Yodish	mean length in standardized version	p-value Wilcoxon
Episode I	42	2.25	1.82	0.001
Episode II	44	2.02	1.89	0.099
Episode III	101	2.24	1.94	< 0.0001
Episode V	89	2	1.98	0.585
Episode VI	29	2.11	2.03	0.365

Table 2: Mean dependency length in Yodish (original and standardized).

Figure 2 displays the frequencies of different lengths in original Yodish and in the standardized version from all five films. The minimum dependency length is 1 – where the head is located immediately before or after the dependent. The maximum was 19, due to a very long sentence. The plot shows that the short dependencies (1 and 2 words) slightly predominate in the standardized version, but the longer ones tend to be more frequent in original Yodish.

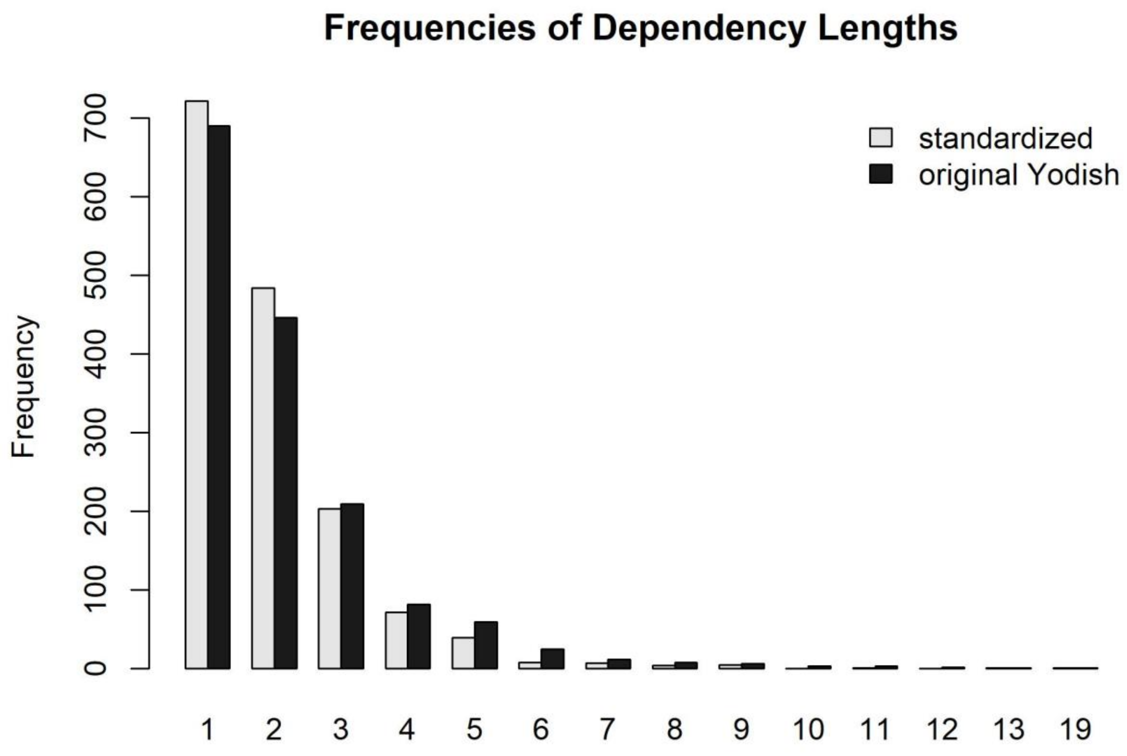

Figure 2: Frequency of dependency lengths in standardized and original Yodish.

In order to interpret these differences, it may be instructive to look at the individual dependencies. The mean lengths were computed for each dependency that occurs 3 times and more. Interestingly, the elements of a nominal phrase have the same mean dependency lengths in the original and standardized versions. These are prepositions, adjectives, nominal modifiers, determiners and numerals. Many dependencies are on average longer in the original Yodish version: auxiliary and modal verbs ('aux'), copulas ('cop'), subject and complement clauses ('csubj' and 'ccomp'), non-finite complements ('xcomp') and adverbial modifiers ('advmod'). Since the auxiliaries and copulas are the most frequent categories among those, the loss of processing efficiency is mostly due to the fact that the auxiliaries and copulas are often separated from the non-finite and nominal parts of the predicates. Below is an example:

(23) a. **Failed** to stop the Sith Lord, I **have**. (E III, original)
b. I **have failed** to stop the Sith Lord. (standardized)

In the original version (23a), the distance is 7 words, while in the standardized version (23b) it is only 1.

At the same time, some dependencies become shorter: passive subjects ('nsubj:pass'), attributive/relative clauses ('acl:relcl') and adverbial clauses ('advcl'). But they are too rare to influence the general picture substantially.

Objects, which were prominent in the previous discussions of Yodish, do not influence the general picture greatly because the move of the object to the beginning of the sentence is compensated by the shorter distance due to the presence of determiners and various NP modifiers (adjectives, etc.) before the object noun. Consider an example:

(24) a. The outlying **systems**, you must **sweep**. (E III, original)
b. You must **sweep** the outlying **systems**. (standardized)

In the original version (24a), the distance between the object *systems* and its head *sweep* is 3 words. In the normalized order (24b), the order is also 3 words. The difference is neutralized due to the presence of the determiner and adjective before the noun.

To summarize, Yodish is less efficient than standard English due to the separation of non-finite, lexical parts from the auxiliaries and copulas. This pattern is not only inefficient, it is also quite unrealistic. The reason is that grammaticalized elements, such as auxiliary verbs, arise in highly predictable contexts. For instance, the future marker *going to* is phonologically reduced (cf. *gonna*) and semantically bleached only in the contexts where it is followed by a verb. When the auxiliary is not accompanied by the lexical part, it is less predictable and therefore less likely to undergo phonological reduction and semantic changes. This frequent co-occurrence of the elements together, which is necessary for grammaticalization, explains why grammaticalized units display formal bondedness with the lexical elements (Lehmann 2015: Section 4.3.2). Moreover, auxiliaries usually lose their positional freedom (*Ibid.*: 168). The existence of auxiliaries in Yodish, which are often split from their lexical elements – which is probably their position in Yoda's native language – is then difficult to explain. This is another piece of evidence showing that Yodish is truly alien.

5 Position of verb in the clause

Most languages of the world place verb in the end of the clause. Ferrer-i-Cancho (2017) argues that this order is efficient because verbs represent cognitive difficulties for language users and learners. For children, verbs are more difficult to learn; actions are harder to verbalize and recall. Non-verbal experiments show a robust strong preference for an order consistent with subject-object-verb even in speakers whose language has a different dominant word order. Therefore, it is efficient to facilitate the processing and the learning of the most difficult item, i.e. verb, by minimizing the uncertainty about the verb. Note that this principle is in competition with dependency length minimization. In particular, SVO languages are also quite common because this order minimizes the distances between the head (the verb) and the dependent arguments.

As was already mentioned, Yodish tends to put the finite verbs (auxiliaries and copulas) last more often than normal English. Is it then more efficient than English because the verbs are more predictable? This is unlikely. It is the lexical verb that should be predicted from the arguments, rather than an auxiliary, which expresses the abstract temporal, aspectual and modal properties of the proposition. However, the lexical verb, as the non-finite part of the predicate, often precedes the arguments. How much does it influence the efficiency of Yodish?

Table 3 displays the frequencies of lexical verbal predicates (no auxiliaries or copulas) in three conditions: a) the verb precedes all other arguments (minimum predictability of the verb); b) one argument (S or O) precedes the verb (medium predictability) and c) two arguments (S and O) precede the verb (maximal predictability). When obtaining the counts, the questions, passives and sentences with clausal complements were omitted. Only the main clauses were analyzed.

While the standardized version clearly prefers medium predictability (due to the default SVO and SV patterns), there are clearly more cases in Yodish both with minimal predictability (58 against 21) and maximal predictability (23 against 0). All examples of the latter type have OSV order:

(25) Grave danger I fear in his training. (E I)
(26) The outlying systems, you must sweep. (E III)
(27) I hear a new apprentice, you have. (E III)

Version	Zero arguments before the lexical verb, minimal predictability	One argument before the lexical verb, medium predictability	Two arguments before the lexical verb, maximal predictability
Original Yodish	58 (37.9%)	72 (47.1%)	23 (15%)
Standardized	21 (13.7%)	132 (86.3%)	0 (0%)

Table 3: Position of lexical predicate in original and standardized Yodish.

In many cases, the final lexical verb is highly frequent and abstract, or 'light', e.g. *have* or *need*, as in (27). There verbs are unlikely to present problems for processing due to their high accessibility. Moreover, the minimal predictability group has the difference of 37 in favour of Yodish, and the maximal predictability group only the difference of 23. This means that Yodish is overall less efficient than English.

6 Discussion

The present study has examined several common tendencies in human languages, which can be explained by considerations of processing efficiency, and investigated whether these tendencies can be found in Yodish.

First, human languages exhibit an implicational relationship between rigid order of subject and object and lack of rich morphology (in particular, case system and agreement). Rigid word order serves as a compensation for lack of morphological tools for expressing 'who did what to whom'. Contrary to this prediction, Yodish has very high variability of the order of subject and object, despite the fact that it has no case marking.

Second, human languages tend to minimize dependency lengths, which saves working memory from overload. Yodish has on average longer syntactic dependencies than the corresponding standard English version. The differences are particularly evident in the prequel films. Longer dependencies mean greater processing load. The main causes of longer dependencies are auxiliaries and copulas, which are separated from the non-finite and nominal predicates. One may wonder how these words have managed to grammaticalize into tense, aspect and modality markers, given that they have no immediate linguistic context to rely on. Their positional variability and lack of bondedness with the lexical part are also highly untypical of grammaticalized units in human languages. This suggests that the grammaticalization processes in Yoda's native language are very different from the ones in human language, if they exist at all.

Third, most languages tend to put verb in the final position, which improves the processing due to its higher predictability given the arguments. However, Yodish more often puts the lexical predicate in the initial position than English, which means that the processing of the verb is not facilitated by the knowledge of the main arguments.

Interestingly, as Yoda's character develops, his word order diverges from standard English. There are differences between the original trilogy and the prequel films. The average dependency length increases. At the same time, the SO and OS entropy decreases, which makes Yodish more consistent and learnable.

How can we explain these differences between human languages and Yodish? First of all, Yoda is a powerful and wise Jedi, who is in command of the Force, which is defined by Obi-Wan Kenobi, another prominent Jedi, as follows:

> ...the Force is what gives a Jedi his power. It's an energy field created by all living things. It surrounds us and penetrates us. It binds the galaxy together (Episode IV).

Thanks to the Force, it is not surprising that the processing capacity of a Jedi is also superior to those of normal humans. At the same time, Yoda does not seem to care much about the fact that his word order may be challenging for normal listeners, since he uses this order not only to speak to other Jedi, but also to normal humans and other species. It is unlikely that he cannot master standard English, or any other language. There are two possible explanations. First, due to his different cognitive make-up, he cannot estimate that this word order is suboptimal for the others. Second, Yoda acts as a dedicated teacher, trying to make his interlocutor's mind work harder, grow and develop.

Finally, let us come back to the real world and take the perspective of George Lucas and other film creators responsible for Yoda's syntax, who seem to be exploiting Levinson's M-heuristic: "What's said in an abnormal way, isn't normal"; or "Marked message indicates marked situation" (Levinson 2000). Extra efforts spent during the processing of Yoda's utterances promise the film audience extra benefits in the form of additional inferences. The story teaches us that even a small green creature with strange English due to immigrant background can be a powerful Master Jedi and a hero. As Yoda says himself, "You must unlearn what you have learned" (Episode V). Thus, although we can conclude that Yoda's word order is suboptimal in the Star Wars universe, it is perfectly efficient for the communication between the film creators and the audience.

References

Barry J. Blake. 2001. *Case*. Cambridge University Press, Cambridge, UK.

Matthew Dryer. 2013. Order of Subject, Object and Verb. In: Matthew S. Dryer and Martin Haspelmath (eds.), The World Atlas of Language Structures Online. Max Planck Institute for Evolutionary Anthropology, Leipzig. (Available online at https://wals.info/chapter/81, Accessed on 2019-04-25.)

Mariya Fedzechkina, Elissa L. Newport and T. Florian Jaeger. 2016. Balancing effort and information transmission during language acquisition: evidence from word order and case marking. *Cognitive Science*, 41(2): 416–446.

Ramon Ferrer-i-Cancho. 2017. The Placement of the Head that Maximizes Predictability. An Information Theoretic Approach. *Glottometrics,* 39: 38–71.

Richard Futrell, Kyle Mahowald and Edward Gibson. 2015. Quantifying word order freedom in dependency corpora. In: *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)*, 91–100, Uppsala, Sweden, August 24–26, 2015.

Edward Gibson. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition,* 68(1):1–76.

John A. Hawkins. 1994. *A Performance Theory of Order and Constituency*. Cambridge University Press, Cambridge, UK.

Roman Jakobson. 1936[1971]. Beitrag zur allgemeinen Kasuslehre. In: Roman Jakobson, *Selected Writings. Vol. II. Word and Language*, 23–71. Mouton, The Hague/Paris.

Paul Kiparsky. 1997. The rise of positional licensing. In: Ans von Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 460–494. Cambridge University Press, Cambridge, UK.

Christian Lehmann. 2015. *Thoughts on Grammaticalization.* 3rd edn. Language Science Press, Berlin.

Stephen Levinson. 2000. *Presumptive Meanings – The Theory of Generalized Conversational Implicature*. MIT Press, Cambridge, MA.

Thomas McFadden. 2003. On morphological case and word-order freedom. In: *Proceedings of the Twenty-Ninth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetic Sources of Phonological Patterns: Synchronic and Diachronic Explanations*, 295-306.

Joakim Nivre, Željko Agić, Lars Ahrenberg et al. 2017. Universal dependencies 2.0 – CoNLL 2017 shared task development and test data. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics, Charles University. http://hdl.handle.net/11234/1-2184. See also http://universaldependencies.org/ (last access 14.12.2017).

Edward Sapir. 1921. *Language, an introduction to the study of speech*. Harcourt, Brace and Co, New York.

Anna Siewierska. 1998. Variation in major constituent order: a global and a European perspective. In: Anna Siwerierska (ed.), *Constituent Order in the Languages of Europe*, 475–552. De Gruyter Mouton, Berlin.

Kaius Sinnemäki. 2008. Complexity trade-offs in core argument marking. In: Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds), *Language Complexity: Typology, Contact, Change*, 67–88. John Benjamins, Amsterdam.

Milan Straka and Jana Straková. 2017. Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe. In: *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, Vancouver, Canada, August 2017.

Russel S. Tomlin. 1986. *Basic Word Order: Functional Principles*. Croom Helm, London.

Jan Wijffels. 2018. udpipe: Tokenization, Parts of Speech Tagging, Lemmatization and Dependency Parsing with the UDPipe NLP Toolkit. R package version 0.7. https://CRAN.R-project.org/package=udpipe