# A NATURAL HISTORY OF $\mu$

Moreno Mitrović

BLED INSTITUTE & UNIVERSITY OF CYPRUS

## Contents

1. **INTRODUCTION**
   1.1 The $\mu$-meanings: a view from Japanese .......................... 4  
   1.2 From conjunction to $\mu$ typologically: a parametric hierarchy . . 6

2. **TOWARD THE COMMON CORE OF $\mu$**
   2.1 A two-tiered alternative system ..................................... 8
   2.2 Spelling out the $\mu$-meanings synchronically ..................... 12
      2.2.1 Polarity sensitivity ............................................ 13
      2.2.2 Universal quantification ....................................... 14
      2.2.3 Freedom of choice ............................................. 15
      2.2.4 Additivity ..................................................... 16
      2.2.5 Type-lifted associates: structurally distinguishing additivity . . 17

3. **TWO HISTORIES OF $\mu$**
   3.1 The Japonic primacy of scalarity ................................. 23
   3.2 The non-priority of polarity sensitivity ............................ 26
      3.2.1 Two changes in Classical Japanese ............................. 27
3.2.2 The rise of polarity sensitivity in Indo-European

3.3 Formalising changes

3.3.1 First approach: the diachronic change in the mode of exhausification

3.3.2 Second approach: dissecting even

3.4 Parametrising the semantic incarnations of μ

4 Conclusions & Outlook
abstract. This chapter reports an inter-genetic diachronic view of the universal superparticle dubbed $\mu$: the logical particle that can, depending on its structural context, behave like a universal quantifier, distributive conjunction, marker of additivity, negative-polarity or free-choice marker. Three language families are investigated, Indo-European, Japonic, and Sino-Tibetan, in order to outline a natural history of $\mu$, i.e., some universal and tangential principles of semantic change.

1 INTRODUCTION

The $\mu$-meanings are those generally characterised by conjunctive meanings or inferences. One of the general concerns of this paper is the question of dissociating language-particular, and random, diachronic developments in the semantics of $\mu$ from the general, and deterministic, diachronic semantic principles that are universal to the $\mu$ formative, regardless of genetic affinity. Since this supposes there is a logical primitive $\mu$ that is common cross-linguistically, I devote a part of this paper to buttressing this view. The argument is two-fold and is reproduced from Mitrović (2014) and Mitrović and Sauerland (2014, 2016). Firstly, if in a single language the construction with a single-meaning $\mu$ can bring about the expression of two meanings, then, ceteris paribus, $\mu$ is hard-wired to express those two meanings. Japanese is used to show this. Secondly, if cross-linguistic expressions of several particular logical meanings are constructed using a single morpheme, without morpho-phonological resemblance or genetic links (say, modern Dravidian and Gothic), then those semantic properties that are detectable from the logical forms of the expressions are attributable to, and evident
of, a common and presumably universal semantic core.

The first hope of this paper is to, therefore, demonstrate that such a core exists. The second desideratum is the detection of some common properties of the semantic change itself that is reflective of the presumably natural principles of language variation and change. The latter is located in the primarily scalar meaning of \( \mu \) in Old Japanese, for which a morphosemantic decomposition is put forth. I will adopt several conditions on interpretation, which are couched in considerations on economy and efficient computation, and state several parameters in this vein.

Before doing, I devote the rest of this section to characterising in informal terms the kinds of \( \mu \)-meanings I investigate (§1.1). I will then couch the \( \mu \)-type expression of conjunction within a wider typology of conjunction in §1.2. In §2 I then motivate a common core for \( \mu \)-meanings and proceed to spelling out those various meanings. §3 represents the diachronic section in which I explicate ‘two histories’ of \( \mu \), that is two diachronic facts from Japonic and Indo-European. This section amount to the development of a working set of parametric properties for \( \mu \) meanings.

1.1 THE \( \mu \)-MEANINGS: A VIEW FROM JAPANESE

Modern Japanese boasts a harmonic system of expression of two classes of logical meanings: those characterised by the \( \kappa \)-superparticle \textit{ka} and those characterised by the \( \mu \)-superparticle \textit{mo}.

This subsection briefly looks at the the ways in which natural language incarnates logical constants such as conjunctive and disjunctive connectives or interrogative, additive and quantificational expressions using a single set of two morphemes. Previous research by Szabolcsi (2010, 2014), Kratzer and
Shimoyama (2002) and Slade (2011), among many others, has established that languages like Japanese may use only two morphemes, *mo* and *ka*, to construct universal/existential as well as conjunctive/disjunctive expressions respectively. Throughout this paper, we abbreviate the Japanese *mo* particle and *mo*-like particles cross-linguistically as $\mu$ and the Japanese *ka* and *ka*-like particles cross-linguistically as $\kappa$.

In Japanese, *mo* also serves as an additive and *ka* as an interrogative element. This semantic multifunctionality of superparticles is clearly exhibited by the following four pairs of examples in (1) and (2), where the left column (1) shows the *mo*-series and the right column (2) shows the *ka*-series.

(1) The $\mu$-series (*mo*)

<table>
<thead>
<tr>
<th>a. ビル(も)メアリーも</th>
<th>Bill <em>mo</em> Mary <em>mu</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>B $\mu$ M $\mu$</td>
<td>‘(both) Bill and Mary.’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b. メアリーも</th>
<th>Mary <em>mu</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>M $\mu$</td>
<td>‘also Mary’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>c. 誰も</th>
<th>who $\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dare <em>mo</em></td>
<td>‘everyone’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>d. どの学生も</th>
<th>dono gakusei <em>mo</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>INDET student $\mu$</td>
<td>‘every student’</td>
</tr>
</tbody>
</table>

(2) The $\kappa$-series (*ka*)

<table>
<thead>
<tr>
<th>a. ビル(か)メアリーか</th>
<th>Bill <em>ka</em> Mary <em>ka</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>B $\kappa$ M $\kappa$</td>
<td>‘(either) Bill or Mary.’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b. 分かるか</th>
<th>wakaru <em>ka</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>understand $\kappa$</td>
<td>‘Do you understand?’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>c. 誰か</th>
<th>who $\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dare <em>ka</em></td>
<td>‘someone’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>d. どの学生か</th>
<th>dono gakusei <em>ka</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>INDET student $\kappa$</td>
<td>‘some students’</td>
</tr>
</tbody>
</table>

When a superparticle like *mo* or *ka* in Japanese combines with two nominal arguments, like *Bill* and *Mary*, coordination obtains, i.e. an expression of conjunction and/or disjunction in presence of the $\mu$ and/or $\kappa$ superpar-
ticle, respectively. When *mo* combines with just one argument (*Mary*), additive (antiexhaustive) expression comes about. When a proposition combines with *ka*, we end up with a polar question (i.e., a set of two propositions). A combination of a superparticle with an indefinite *wh*-expression, like *dare* ‘who’ (10–22), delivers a quantificational expression, either with an existential flavour (‘*someone*’, *dare-ka*) or a universal flavour (‘*everyone*’, *dare-mo*). Similarly, non-simplex quantificational expressions like ‘*some/every student/s*’ obtain in Japanese when an indeterminate *wh*-phrase, like *dono*, combines with a nominal like ‘*student(s)*’.

We assume that the two series of superparticle meanings in (1) and (2) do not result from homophony, contra Hagstrom (1998) and Cable (2010), as argued by Slade (2011), Szabolcsi (2013), and Mitrović and Sauerland (2014, 2016).

1.2 FROM CONJUNCTION TO \( \mu \) TYPOLOGICALLY: A PARAMETRIC HIERARCHY

The aim of this subsection to parametrically contextualise the availability and general place of \( \mu \) within a larger system of cross-linguistic expression of conjunction. Since 67% of world languages express \( \mu \) meanings, i.e. employ conjunction markers which also express non-conjunctive meanings (Gil, 2005), I address, and programmatically propose, a typological system of parametric hierarchies in the spirit of Biberauer and Roberts (2017).

While most languages of the world have some flavour of \( \mu \) (Gil, 2005), not

---

1 A combination of a *wh*-term with \( \mu \) is, *prima facie*, ambiguous between a universal distributive and a polar indefinite expression. Prosodic cues to disambiguation have been proposed: see Szabolcsi (2010, 202), Nishigauchi (1990), Yatsushiro (2002), Shimoyama (2006, 2007), among others, for an account of the synchronic distribution of facts. I show in the following sections that the universal distributive semantics of *wh*-\( \mu \) is diachronically primary in the history of Japonic and developed a diachronic analysis of the the rise of polarity sensitivity.

2 See Shimoyama (2007) for an elegant and convincing analysis.
all do. I outline the prerequisite ingredients for a language to lexicalise μ here.

(3) A parametric hierarchy for coordination:

```
<table>
<thead>
<tr>
<th>Is coordination lexicalised?</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
</tr>
<tr>
<td>YES</td>
</tr>
</tbody>
</table>
```

- Middle Egyptian (WALS=6)

```
<table>
<thead>
<tr>
<th>Is the ∧/∨ logical contrast lexicalised?</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
</tr>
<tr>
<td>YES</td>
</tr>
</tbody>
</table>
```

- Warlpiri
- ASL

```
<table>
<thead>
<tr>
<th>Is conj. sensitive to N/V categorial contrast?</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
</tr>
<tr>
<td>YES</td>
</tr>
</tbody>
</table>
```

- Universal
- Unattested

```
<table>
<thead>
<tr>
<th>Is disj. sensitive to N/V categorial contrast?</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
</tr>
<tr>
<td>YES</td>
</tr>
</tbody>
</table>
```

- V-type/C-level conj. employed for conj of all categories.
- [most modern IE languages, …]
- [67% of langs.]
- N-conj. marker
- the μ superparticle

The main question I am concerned with in this section is the semantic switch between conjunction generally and the expressibility of μ meanings. The predictive power of the parametric hierarchy is implicational: a language that does not lexicalise the semantic contrast between conjunction and disjunction is, logically, predicted to be able to lexicalise μ meanings associated with nominal conjunction and quantificational uses.

Another conjectured universal regards the categorial sensitivity of coordination. While conjunction has been shown empirically not to be syncategoric: Zhang (2010); Mitrovič and Sauerland (2016), disjunction appears to be different in this regard. I am unaware of a language which could em-
ploy a different disjunction morpheme for distinct categories.

2 TOWARD THE COMMON CORE OF $\mu$

The proposed common core of $\mu$ meanings, valid and operative synchronically, is related to alternatives and the exhaustification process.

The core motivation for this is the ability for the proposed analysis to account, in a unified way, for additivity and negative polarity. The exhaustification operator, $\text{exh}$, deliver negative polarity meaning if applying once and additivity if applying twice. I am unaware of any other system that can deliver the two types of meanings by regulating, in a principled way (as I discuss briefly below), the iteration parameter alone. On a general level, this is a desirable economy consideration of the system.

2.1 A TWO-TIERED ALTERNATIVE SYSTEM

I briefly expound here on the formal system and the assumptions I use in the paper.

The account I develop rests on the assumption that pragmatic devices for reasoning are anchored in narrow syntax in the sense that Chierchia (2013) develops in some detail. Such a device is a covert exhaustification operator, $\text{exh}$, which is able to derive polar-sensitive, free-choice, or additive meanings as well as scalar implicatures (SI). This subsection introduces Chierchia’s (2013) system using an example from disjunction.

As a starting exemplar of SIs, we take disjunction which, in English at least, carries an obligatory implicature, as Chierchia (2013: 429) notes: ei-

Benjamin Slade (pers. comm.) brings the Kannada facts to my attention (Amritavalli, 2003); the contrast between -oo and illa seems to be more complex with regards to the non-distjunctive meanings of illa and does not seem to be a categorial one. I leave this question, as well those relating specifically to $\kappa$-meanings, aside in this paper.
ther an epistemic one, where a disjunctive expression $p \lor q$ implicates speaker’s ignorance since the speaker doesn’t know whether $p \lor q$; this yields epistemic possibilities that $\diamond p$, $\diamond q$, $\diamond [p \lor q]$ or $\diamond [p \land q]$ as per (4a). We do not concern ourselves with ignorant readings of disjunctive sentences here. What we do concern ourselves with is the other possible, and opposite (i.e., counter-ignorant), implicature that disjunction generates, namely the scalar implicature (SI) where $p \lor q$ is enriched by denying conjunction so as to mean $[p \lor q] \land \neg[p \land q]$ as sketched in (4b).

(4) Mary saw John or Bill.

a. ignorance implicature

i. $\diamond [j] \land \diamond [b] \land \diamond [j \lor b] \land \diamond [j \land b]$

ii. ‘The speaker doesn’t know whether Mary saw John and the speaker doesn’t know whether Mary saw Bill and the speaker doesn’t know whether Mary saw John and Bill.’

b. scalar implicature

i. $[j \lor b] \land \neg[j \land b]$

ii. ‘Mary saw John or Bill but not both.’

The explicans for SIs we adopt rests on the process of exhaustification. More specifically, the preliminary framework within which we locate our analysis is that of Chierchia (2013), int. al., who locates the aetiology of SIs, an inherently pragmatic phenomenon, in a grammatical, or more precisely, a narrow-syntactic, module. This is achieved, in simple terms, by positing a covert exhaustification operator ($\text{exh}$) in the narrow syntax which attaches to the root of some proposition-level syntactic structure, say IP (for Chierchia 2013) or CP (for Mitrovic 2014), and post-syntactically exhaustifies
the proposition against a particular set of alternatives pre-determined in an Agree-wise fashion in the syntax. The lexical entry for this exhaustification operator, of type \(\langle\langle(s,t)\rangle(t,s)\rangle\) and labelled \texttt{exh} here, is given in \ref{c1} and reads in informal language as bearing the following meaning: the assertion, \(p\), is true and any non-entailed alternative to the assertion, \(q\) an alternative, is false.

\[
(5) \quad \texttt{exh}(p) = p \land \forall q \in \mathfrak{A}(p)\left[ [p \vdash q] \rightarrow \neg q \right]
\]

We recognise two kinds of alternative sets that may be generated. One is the set of sub-domain (\(\delta\)) alternatives, which excludes any Boolean or (strictly) scalar terms. This set of \(\delta\)-alternatives is assumed to be particularly relevant in Focus calculation (cf. \cite{FoxKatzir2011}). The other alternative kind comprises a (strictly) scalar set of alternatives (\(\sigma\)), which includes only the scalar alternatives of a given expression. Scalar terms like \textit{all}, \textit{some}, \textit{and or or} are just some of scalar terms with only scalar alternatives (pace \cite{Sauerland2004}, \textit{int. al.}, but see \ref{??}). Thus an \texttt{exh}-operator specified with a \(\sigma\)-feature will post-syntactically (pragmatically) exhaustify a given proposition (syntactically, its sister) against the scalar alternatives to that proposition (and only the scalar alternatives). A \(\delta\)-carrying \texttt{exh} will exhaustify a proposition and deny the sub-domain alternatives, and only those alternatives, to that proposition. The featural specification on \texttt{exh} are relegated to syntactic rules on Agree, via which the range of exhaustification is determined. The entire alternative set to a disjunctive proposition like \ref{c2}, thus takes a two-dimensional shape consisting of sub-domain \(\delta\)- and strictly scalar \(\sigma\)-alternatives, as sketched in \ref{c3}. The sub-domain (\(\delta\)) alternatives are

\[Caching\: in\: the\: presupposition\: of\: p,\: then\: \texttt{exh}(p) = \forall q \in \mathfrak{A}(p)\left[ [p \vdash q] \rightarrow \neg q \right]\]. We will also, in \ref{??}, modify the LF of \texttt{exh} minimally with reference to Innocent Exclusion.
plotted horizontally and comprise two singleton propositions: \( j \) for ‘Mary saw John’ and \( b \) standing in for ‘Mary saw Bill’. The vertical axis features the two scalar alternatives: \( j \land b \) which reads ‘Mary saw John and Mary saw Bill’ and \( j \lor b \) which is a shorthand for ‘Mary saw John or Mary saw Bill’ (the assertion).

\[
(6) \quad \mathfrak{A}(\mathfrak{P}) = j \quad \begin{array}{c}
\text{assertion} \\
j \lor b
\end{array} \quad \begin{array}{c}
\delta\text{-alternatives (}\delta\mathfrak{A}\text{)} \\
j \land b
\end{array} \quad \begin{array}{c}
\sigma\text{-alternatives (}\sigma\mathfrak{A}\text{)}
\end{array}
\]

Note also that the proposed two-tiered system of alternative structure is also analogous to the lattice-theoretic notions of chains and antichains. 5

\[
(7) \quad \text{i. } D_\delta = \delta = \textbf{Antichain}(\mathfrak{A}) = \mathbf{A}(\mathfrak{A}). \text{ No two elements are comparable. } \forall x, y \in \delta, \text{ neither } x \leq y \text{ nor } y \leq x \text{ hold. The } \mathbf{A}(\mathfrak{A}(\mathfrak{P})) = \{j, b\}
\]

\[
\text{ii. } D_\sigma = \sigma = \textbf{Chain}(\mathfrak{A}) = \mathbf{C}(\mathfrak{A}). \text{ Any two elements are comparable. } \forall x, y \in \delta, \text{ either } x \leq y \text{ or } y \leq x \text{ hold. The } \mathbf{C}(\mathfrak{A}(\mathfrak{P})) = \{j \lor b, j \land b\}
\]

Enriched (exclusive) disjunction may be derived through both means of exhaustification: either globally through \texttt{exh} attaching to the disjunctive phrase and triggering scalar exhaustification, where the effect of (4b) is derived, as per (8a). Alternatively, exclusive disjunction is also derived locally via \texttt{exh}-attachment to each of the (two) disjuncts and exhaustifying them with respect to their respective \( \delta \)-alternative set in a Focus-like fashion, as shown in (8b). The two exhaustification strategies yielding the exclusive implicatures are calculated by negating different alternative (sub-) sets as per (8a-ii) and (8b-ii).

5 This will become relevant and useful in §3.
Two ways of calculating the SI of (4) and deriving the exclusive component:

a. **GLOBAL CALCULATION** of the exclusive component via $\text{EXH}_{[\sigma \mathcal{A}]}$

   i. Syntactic structure (simplified):

   $$
   \text{EXH}_{[\sigma \mathcal{A}]} \rightarrow j \rightarrow \text{OR}_{[\sigma \mathcal{A}]} \rightarrow b
   $$

   ii. Logical form:

   $$
   \text{EXH}_{[\sigma \mathcal{A}]}(j \vee b) = [j \vee b] \land \neg[j \land b]
   $$

b. **LOCAL CALCULATION** of the exclusivity component via $\text{EXH}_{[\delta \mathcal{A}]}$

   i. Syntactic structure (simplified):

   $$
   \text{EXH}_{[\delta \mathcal{A}]} \rightarrow j[\delta] \rightarrow \text{OR} \rightarrow \text{EXH}_{[\delta \mathcal{A}]} \rightarrow b[\delta]
   $$

   ii. Logical form:

   $$
   \text{EXH}_{[\delta \mathcal{A}]}(j \vee b) = \text{EXH}(j) \lor \text{EXH}(b) \vdash \neg[j \land b]
   $$

These two exhaustification strategies are equally reasonable hypotheses for deriving enriched disjunction in English. The disjunction morpheme $[\text{or}]$, under very natural assumptions pervading the field, directly maps onto the denotation of ‘$\lor$’, a logical disjunction in itself a scalar-alternative-sensitive operator that is targeted by a silent (probing) $\text{EXH}$.

### 2.2 SPELLING OUT THE $\mu$-MEANINGS SYNCHRONICALLY

With the basic alternative-based technology in place, and the assumption that the core of $\mu$-meanings is the alternative-triggering signature, I now
proceed to spelling out how various meanings arise as an interplay of various kinds of alternatives at play and various modes of exhaustification.

The toy sample of deriving exclusive disjunction was appropriate for the purposes here since I assume that wh-phrases denote existential quantification (Xiang, 2016), which is homeomorphic to discrete disjunction.

### 2.2.1 Polarity Sensitivity

My proposal which assumes iterative exhaustification as the device that delivers FCIs, additives, and universals, is based on an economy consideration. \( \text{exh} \) will iterate only when it has to, everything else being equal. Crucially, NPIs, as \( \mu \)-marked wh-indefinites in the scope of negation provide non-trivial contexts in which a single mode of exhaustification is not only successfully meaningful but also economically preferred. This is in line with Chierchia’s (2013) Blocking condition (where ‘>’ is an economy preference symbol), given in (9).

\[ \text{exh}_6[ \text{neg} \cdots \text{DP}_\delta \cdots ] > \text{exh}_δ^* [ \text{neg} \cdots \text{DP}_\delta \cdots ] \]

Assume now a historical change where the ‘last resort’ recursive exhaustification (6) become a default ‘first response’ mode of interpretation. Phase II languages (??), such as Gothic or Modern Greek, encode anti-exhaustivity as a ‘first response’ mode. Such languages are predicted not to be able to encode polarity sensitivity, which requires a single layer of exhaustification.

Given the distribution to Downward Entailing contexts, the iterativity as a rescue operation is blocked. In what follows, I advocate another presupposi-

---

6 The blocking condition is that of Chierchia (2013: 278, 57) but the notion of pre-exhaustivised alternatives is encoded, as an instance of recursive sub-domain exhaustification (exh\(_δ^*\)), as discussed above.
tional condition which distinguishes NPIs from other meanings.

There is another property that distinguishes NPIs from other meanings. The difference between $\mu$-marked NPIs and other $\mu$-marked indeterminate ($wh$-phrase) expressions can be derived using a semantic and a syntactic parameter. The former is Chierchia’s (2013) ‘Proper Strengthening’ (PS) parameter (10). The latter is a featural requirement on $\mu$ for there to be restricted to Downward Entailing contexts (which can be achieved using a Minimalist checking theory). In this paper, I focus on the semantic parameter.

(10) **THE PROPER STRENGTHENING PARAMETER** (Chierchia, 2013: 274, 48)

\[
\text{exh}^{PS}_\lambda(\phi) = \begin{cases} 
\text{exh}_\lambda(\phi) & \text{if } [\text{exh}_\lambda(\phi) \sqsubseteq \phi] \\
\bot & \text{otherwise}
\end{cases}
\]

Since $\text{exh}^{PS}$ is a presuppositional operator, this allows us to capture the parameter: either $\text{exh}$ is presuppositional, in the sense of PS (10), or it is not. If it is, $\text{exh}$ (recursive or not) is required to yield a meaning which is informationally stronger than the assertion. If it is not, non-stronger meanings pass. This derives the distribution of NPIs.

2.2.2 **UNIVERSAL QUANTIFICATION**

Another $\mu$-meaning is that of universal quantification, as instantiated in Gothic (structurally in the C-domain) or Hittite (structurally in the nominal domain). What captures the meaning of a universal quantifier is a iterative sub-domain ($\delta$-level) exhaustification. Note that this way of obtaining uni-

---

As well as various kinds of FCIs contrasting in quantificational force, such as the German *irgend*-type versus the Italian *qualunque*-type of FCIs, as Chierchia (2013: 275ff.) discusses in detail.
versal quantification via iterative application of \textit{exh} does not violate Chierchia’s (2013) naturally defined exhaustification economy:

(11) \textbf{EXHAUSTIFICATION ECONOMY} \textit{[Chierchia, 2013: 129, ex. 75]}

Avoid unnecessary exhaustification.

a. \(\star \text{EXH}_A [Y \cdots X_{[+A]} \cdots],\) if the result is itself a member of \([[Y]_A]\) different from \([X]\).

b. \(\star \text{EXH}_A [Y \cdots X_{[-A]} \cdots],\) if logically equivalent to \([Y \cdots X_{[-A]} \cdots].\)

The way in which the proposed derivation of universals in Gothic or Hittite proceeds abides by the economy principle in (??), despite the over-generation problem of FC-effect that Chierchia (2013: 122, fn. 30) warrants against. Recursive exhaustification over the domain of non-scalar alternatives does not violate (??), as shown in (11), taken from Chierchia.

(12) \textit{exh}^2([p \lor q]) = [p \lor q] \land \neg\text{exh}(p) \land \neg\text{exh}(q)

\quad = [p \lor q] \land \neg(p \land \neg q) \land \neg(q \land \neg p)

\quad = [p \lor q] \land (p \rightarrow q) \land (q \rightarrow p)

\quad = p \land q

2.2.3 FREEDOM OF CHOICE

Universal FCIs require wide-scope (more on that below) and, while they derive via iterative sub-domain exhaustification just like distributive universal quantifiers, they have additional characteristic, namely being constrained by Fluctuation (13). I take Fluctuation to be a privative character of (universal) FCIs.

Therefore, FCIs share with universal quantifiers the iterative mode of being derived, to the exclusion of NPIs (cf. 9), which also do not satisfy the PS
condition \[10\]. The main property that discriminates between the two types of meanings is, informally, licensing restriction to modalised contexts. In this regard, I propose that the discriminating factor is the Fluctuation Constraint [Dayal (2009)].

\[13\] **THE FLUCTUATION CONSTRAINT** [Dayal, 2009, 241, 5b]

\[
\neg \exists x \forall w'[w_a \leq w'] \lambda x [P(w')(x) \land Q(w')(x)] = X
\]

The last meaning I sketch is that of standard additivity, which shares technical character with related meanings but differs in more morphosyntactically detectable aspects.

**2.2.4 ADDITIVITY**

How are additive meanings technically different from other \(\mu\)-meanings? I now turn to explicating the commonalities and differences. The main difference can be stated in purely syntactic terms: \(\mu\)-encoded meanings of universals, FCIs, and NPIs arise when the \(\mu\)-associate is a \(wh\)-phrase. Additives arise as an elsewhere option on \(\mu\)-associating (complementation).

This same distribution can also be captured semantically by cashing in on the existential quantifier meaning of \(wh\)-phrases, which is what I assume (see Xiang (2016) and those she cites for evidence). A proper noun like ‘John’ has no hardwired existential meaning: therefore, when \(\mu\) associates with ‘John’ it will trigger its alternatives and recursive exhaustification takes place (see Mitrović and Sauerland 2016 for details) to yield an antie exhaustive, or additive, meaning. Additive expressions of \(wh\)-phrases are therefore predicted to be impossible \(\mu\)-meanings.

---

8 See also the discussion in Chierchia (2013: 31ff.), who provides a modified definition of Fluctuation; his variant does not have bearing on my purposes here.
The differentiating property of additives is therefore the restriction to non-
wh-associates, which can be modelled semantically as a possible restriction
on type-association (e-type versus types higher than e), or a restriction on
association with existential quantifiers (assuming wh-phrases are, indeed,
existential quantifiers). I adopt a lighter take here, assuming that the dif-
ferentiating property is that of ‘indeterminacy’ (in the sense of Shimoyama
2006: while wh-phrases are indeterminate, proper DPs, such as ‘John’ are
not.

Before turning to a diachronic analysis in § 3, I discuss in the remainder
of this section the commonalities and differences between obligatorily type-
lifted wh-phrasal μ-associates and non-wh-associates.

2.2.5 TYPE-LIFTED ASSOCIATES: STRUCTURALLY DISTINGUISHING
ADDITIVITY

In Indo-European, all archaic languages operating a μ particle show type-
lifted associates. In Hittite, as Hoffner and Melchert (2008: 149) note, the
universal distributive expressions like ‘each(one), every(one)’ quantificational
expressions correspond to kuišša (𒆪𒄿𒋾𒈬), comprising of an inflected wh-comp-
onent kui- (𒆪𒄿𒈬) and the conjunctive (super)particle -a (𒈬) / -ya (𒈬𒈬), corre-
sponding to our μ category. The following two examples show μP universals
as animate (‘who’) subjects and inanimate (‘what’) objects. The relevant
μ-phrases are bracketed.

(15) nu [dušum.meš-šu kuišš- a] kuwatta utnē paizzi
J sons.his who-μ = V somewhere country.loc went
‘Each of his sons went somewhere to a country.’ (KBo. 3.1.17–18)

(16) nu [kuiṭ-a] arhayān kinaizz[i] what-μ = ∀ seperately sifts

‘She sifts everything seperately.’ (KUB XXIV.11.III.18)

Given the proposed character of μ, I now derive the interpolation of the wh-element in (15). Following Mitrović and Sauerland (2016), I assume is an e-type operator: in (15), in the absence of the wh-pronoun kuišš- ‘who’, additivity is predicted to obtain (‘also his sons’) instead of a universal quantificational term (‘each of his sons’). In line with Mitrović and Sauerland (2016), for the expression universal quantification (15), I follow Shimoyama (2006) to assume that the wh-pronoun in (15) is indeterminate. As such, it is interpreted as a set of type ⟨e, t⟩ restricted to humans, i.e. [kuišš-] = λx.x is human. For the quantificational expressions such as ‘dumu.meš-šu kuišš-a’ (‘each of his sons’), I analyse the μ-associate ‘dumu.meš-šu kuišš-a’ to be interpreted as a set of type ⟨e, t⟩: the set of his sons.

Conversely, ‘his sons’, lacking the wh-pronimal type-lifter, would yield an additive meaning when associating with μ bare.

We maintain the anti-exhaustive semantics for μ particles and account for the universality by specifying that the anti-exhaustive operator (exh R or ¬exh) exhaust subdomain alternatives alone, ignoring the scalar alternatives. This computational instruction is easily hardwired into the syntax by specifying the operator with a [δ]-feature. This line of reasoning is also explored by Chierchia (2006) and, most recently, in Chierchia (2013: 311, ex. 18) who works with pre-exhaustified alternatives, which he encodes on the exhaustivity operator by specifying it with a [Exh-δR]-feature, signalling that the alternatives being exhaustified have already been pre-exhaustified.
While this delivers the same anti-exhaustivity result, note that under our analysis, we are not dealing with pre-exhaustified alternatives but simply allow for recursive (re)exhaustification. The iterative modes of running the Gricean reasoning are technically different but essentially the same (Chierchia 2013: 119 notes this too). Since nothing hinges on this, we the notions of pre-exhaustified alternatives and recursively exhaustified alternatives to be on a par and, for all relevant matters, equivalent (as noted in (18b-ii), too).

Take the syntax of (16) as a working example. Without the μ particle -a, the object would be a wh-DP with indefinite or existential meaning. Since, as Chierchia (2013: 357) eloquently puts it, “indefinites of all colouring receive meanings identical to those of or ... as they are just potentially infinite disjunctions,” (16) sans the meaning contributed by μ would be something along the lines of (17), where ‘sand’, ‘flour’, and ‘ash’ in (17b) are some possible contextual extensions of some things that Hittites may have sifted. We adopt a shorthand for these discrete disjunctions at propositional level in (17c), where ‘she sifts sand’ is abridged as a, etc.

(17) \[ (16) \backslash [\mu] = \]

a. she sifts something

b. she sifts sand \textcircled{\textbf{a}} \textcircled{\textbf{b}} she sifts flour \textcircled{\textbf{a}} \textcircled{\textbf{b}} ...

c. \textcircled{\textbf{a}} \textcircled{\textbf{b}}

The presence of the μ particle thus entails two procedures: (i) the activation of δ-alternatives and (ii) their recursive exhaustification. Still working with the sifting example in Hittite (where we represent only a and b as working alternatives), we arrive at the following alternative schemata and ex-
haustification, which delivers the correct computations.

\[(18)\]  
\[a. \text{ Active } \delta\text{-alternatives:}\]
\[
\begin{array}{c}
\text{assertion} \\
\hline
a \lor b \\
\hline
a \\
\hline
b \\
\hline
\delta\text{-alternatives}
\end{array}
\]

\[b. \text{ Exhaustification:}\]

\[i. \text{ First level:}\]
\[
\text{exh}_{\delta A}(a \lor b) = \text{exh}(a) \land \text{exh}(b) = \bot
\]

\[ii. \text{ Second level:}\]
\[
\text{exh}_{R\delta A}(a \lor b) = \text{exh}_{\text{exh-}\delta A}(a \lor b)
\]
\[
= \neg \text{exh}(a) \land \neg \text{exh}(b) = a \land b \neq \bot
\]

The resulting conjunction is equivalent to universal quantification which is in line with the meaning in (16) we set out to compute. Note that (18b-ii) also shows the iterative requirement of exh and the way in which we restrict the iterative application of exh. In the case of (18b-ii), exh must apply iteratively, that is twice, since a single level of exhaustification leads to a contradiction. The second layer of exhaustification, however, is no longer contradictory (\(\text{exh}_{R\delta A}(a \lor b) = \neg \text{exh}(a) \land \neg \text{exh}(b) = a \land b \neq \bot\)).

Under the assumption of obligatory iterativity of exhaustification, we can thus derive the universal distributive inference using second-level exhaustification yielding an anti-exhaustive inference. Such recursive exhausification approaches, generally stemming from Fox (2007), have recently been advocated widely, for instance in Bowler (2014), Mitrović (2014), Mitrović and Sauerland (2014, 2016), Bar-Lev and Margulis (2014), Singh et al. (2016), and most recently in Szabolcsi (2017a, b). We have, however, departed from Fox (2007) and Chierchia (2013) in that we are recursively exhaustifying over...
a scalar set of alternatives. In sum, the two means of deriving the universal inference, using recursive exhaustification targeting (a) subdomain and/or (b) scalar alternatives, are given in (19).

(19) Recursive exhaustification of ...

a. $\delta$-alternatives:

\[
\begin{array}{c}
\forall \\
\land a \land b \land \\
\land c \land \ldots \\
\land \neg \text{EXH}(a) \land \neg \text{EXH}(b) \land \\
\land \neg \text{EXH}(c) \land \ldots \\
\end{array}
\]

b. $\sigma$-alternatives:
In the next section, I investigate the diachronic processes that gave rise to the synchronic distribution of μ-meanings I have just overviewed.

3 TWO HISTORIES OF μ

There are two diachronic facts underlying the meanings of μ. In Japonic, μ originates as an exclusively scalar particle. In both Japonic and Indo-European, μ-marked NPIs are secondary and develop from universal quantificational or universal free-choice expressions.

I first take each of the two facts in turn in subsections §3.1–3.2.
3.1 The Japonic Primacy of Scalarity

The oldest text in Japonic dates back to the 8th century AD and allows us to see how the contemporary superparticle system developed. Unlike the latter, Old Japanese (OJ) superparticle mo did not have the role of performing conjunction, nor encoding negative polarity.

What is more, OJ μ shows that the stipulated [σ] feature and the general scalar dimension of meaning is empirically motivated. I will first introduce the data in Old Japanese before addressing the changes that occurred in the Classical Japanese period in §3.2.1.

In the earliest OJ corpus (Man’yōshū MYS, 8th c.), the [wh+μ] quantificational expressions were confined to inherently scalar (σ) complements, as first noticed by Whitman (2010).

Not only was the polar construction absent from the μ-system in the OJ corpus, but μ₀ subcategorised for scalar hosts only. That is, the only eligible hosts of mo were either numeral nominals or inherently scalar wh-terms: how-many and when. The combination of a numeral π and μ, yielded the least likelihood reading along the lines of ‘even π’. In the wh-domain, the μ particle created universal quantificational expressions as shown below.

Chierchia’s (2013) system gives us the descriptive power to label this μ as carrying [uσ] since non-scalar complements were disallowed:

(20) 以都母 々々々 於母加 古比 須々
    itu-mo itu-mo omo-ga kwopi susu
    when-μ when-μ mother-GEN yearning by

    ‘I always, always think of my mother [i.e. at all times]’

(MYS, 20.4386; trans. by Vovin 2013: 146)

The details of the system presented in Mitrovic (2014) in fact predict the diachronic behaviour according to which the rise of conjunction or additivity is precluded in absence of the δ-feature on μ.
As there have been few nights in which we slept together …

(MYS 5.804a, ll. 46–47)

To buttress the fact that only scalar wh-terms were allowed, see Table 1 where counts of μ hosts are given.

<table>
<thead>
<tr>
<th></th>
<th># of attestations</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCALAR [wh+μ]</td>
<td>total 24</td>
</tr>
<tr>
<td>itu mo ‘when μ’</td>
<td>12</td>
</tr>
<tr>
<td>iku mo ‘how much/many μ’</td>
<td>11</td>
</tr>
<tr>
<td>NON-SCALAR [wh+μ]</td>
<td>total 0</td>
</tr>
<tr>
<td>ado/na/nado mo ‘what/why μ’</td>
<td>0</td>
</tr>
<tr>
<td>ika mo ‘how μ’</td>
<td>0</td>
</tr>
<tr>
<td>ta mo ‘who μ’</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: Distribution of ± scalar μ-hosts in OJ

Let me now turn to showing the scalar distribution of the OJ μ particle in non-wh-constructions. In (22), the clause-final verb constitutes a downward entailing context.

(22) 吾 懲 流 千 重 乃 一 重 μ
(w)a ga kwopuru ti pye no pito pye mo
PRO.1 GEN love.ADNOM 1000.NUM layer/part GEN person one μ
人 不令知
pito sire-zu
person tell-NEG.AUX

‘Without a thousandth part of my great love’ (万葉集, Man’yōshū, 13, 3272, ll. 25–27, trans. Faki et al. 2005: 309, #983-4, l. 14)

(23) [久 爾 能 富 μ] 美 由
kuni no po mo mi yu
land/area GEN fire/highest μ see PASS

‘the highest part of the land is visible’ (KK., 41 l. 6, trans. by Philippi, cf. NSK. 34)

10 In the examples, # signals a line break in the poems.
Even the highest parts?

(24) [[阿佐 妝 耳 毛] # 伊弟低 由介那]
asatwo nimothetoyo go.out.

‘Even its morning-door # We would go forth from—’ (NSK, 16, ll. 3–4, transl. by Aston 1896)

In (25), the scalar component is visible through the use of the restrictive particle dani, which combines with the nominal ‘rice’. In association with μ, the scalar additive meaning of ‘even’ obtains:

(25) 梨梅 多僕 母 # 多礙底
kome mo tagete
rice.deverbal restr.ptc μ let.eat/drink.grnd

‘Pass on—having stolen # Even the very rice,’ (NSK, 107, ll. 3–4, trans. by Aston 1896)

Example (26) shows a rare conjunction-like expression of μ where the relevant (negative) scale that μ associates with is that of ‘(not) knowing’ on part of the speaker.

(26) 比騰 能 # 於謀提 母 始羅孺 # 伊弊 母 始羅孺
pito nomote mo sira-zu ipye mo sira-zu
person gen front μ know-NEG.aux house μ know-NEG.aux

μ

‘I know not the face, # Nor do I even know the house’ (NSK. 111, ll. 4–5, trans. by Aston 1896)

The OJ mo, given its restriction to scalar complements, carries an uninterpretable σ-features which sufficiently captures its distribution.

While assuming a Minimalist system of syntax, I assume, in line with Roberts (2010), that an
(27) OJ: $\mu[uo]$

Furthermore, this paper argues for the following characterisation of $\mu$-meanings in OJ.

(28) In Old Japanese $\llbracket \mu \rrbracket(x)$ is defined iff $x$ denotes a a totally ordered set, chain $C$, morphosyntactically encoded with the $[\sigma]$ feature. $\mu(x)$ denotes the maximal member of that chain, given as an application of the ceiling function of that chain, $\lceil C \rceil$.

The fact that scalar $\mu$-expressions in OJ under negation resist those negative inferences that obtain in NPIs, they may be analysed as Positive Polarity Items (PPIs). However, as Szabolcsi (2004: 419) notes, “Some-type PPIs do not occur within the immediate scope of a clausemate antiadditive operator.” As seen in (21), ｙwoo-no ikuda mo occurs within the scope of clausemate negation ($ne$-). More crucially, the term in non-negative contexts has a clearly universal meaning which weakens under negation but does not leave the positive scale. Thus, such constructions are best analysed as SIs. In (29), an informal sketch of positive/negative inferences is given based on the example (21).

(29) $\llbracket [\text{not [all nights]}] \rrbracket = \begin{cases} \rightsquigarrow \text{some nights} & \text{(scalar reading)} \\ \rightsquigarrow \text{no nights} & \text{(polar reading)} \end{cases}$

3.2 THE NON-PRIORITY OF POLARITY SENSITIVITY

Within the range of meanings encoded by the Mu superparticle, the polarity-sensitive encoding has a diachronic aetiology in both IE and JP. Methodolog-
ically, at least, I claim that the identity of diachronic trends is evident of a (conjectured) universal principle:

(30) $\mu$-encoding of polarity sensitivity is not a primary incarnation of $\mu$-meaning. These originate as universals and are diachronically reanalysed as weakened under negation.

Let me first show the rise of polarity sensitivity in Japonic (§3.2.1) before turning to IE (§3.2.2).

3.2.1 TWO CHANGES IN CLASSICAL JAPANESE

Two interlocked changes can be detected in the classical period. We take each of them in turn.

The first change concerns the loss of restriction on the type of complements that the Classical Japanese $\mu$ may associate with. Unlike in the Old Japanese period, Classical Japanese $\text{mo}$ can be seen to freely associate with hosts of non-scalar type. One such example is in (31), where $\text{mo}$ associates with a $\text{tare}$ ‘who’ which has as a restriction set the non-scalar subdomain of all individuals.

(31) たれも 見おぼさん 事
$t\text{are mo}$ mi-obos-an koto
who $\mu$ see.INF-think.HON-TENT/ATTR matter
‘the fact that everybody wanted to see’ (HM II:226/2; Vovin 2003: 128)

Given the system we propose, the Classical Japanese $\mu$ is parametrised as having the $[\delta]$ feature both present and, in the case of (31), set to a positive setting. Recall that Chierchia’s (2013) systems of morphologically marked PSI requires that at least one of the $[\delta, \sigma]$ features be positively set.
Mitrović’s (2013) system, in fact, predicts that if both $[\sigma]$ and $[\delta]$ are available features polarity sensitivity should obtain by definition of the system, *ceteris paribus*. As it happens, this is exactly what we find in non-archaic Japonic.

The rise of $\mu$-marked polarity sensitivity is evident from the data in (32) where a *wh*-pronoun *nani* ‘what’ with a non-scalar domain extension is interpreted under negation.

(32) いまはなのに心もなし
ima fa nani-no kokoro mo na-si
now top what-gen idea $\mu$ NEG-FIN

‘I do not have any thoughts [but of meeting you] now’

(IM XCVI: 168.9; Vovin 2003: 424)

Therefore, the Classical (early middle) Japanese $\mu$-system can be analysed as having lost the licensing restriction to $\sigma$-contexts. This can be represented by assuming that, unlike OJ, both $[\sigma]$ and $[\delta]$ features are interpretable on $\mu$, or that neither are (i.e., there is no narrow syntactic restriction posed on the structural context, or complementation).

(33) OJ: $\mu[\langle u \rangle \sigma] \xrightarrow{\text{cca. 250 years}}$ CJ:

$$\begin{align*}
\mu_1 & \begin{bmatrix} \langle u \rangle \sigma \end{bmatrix} \\
\mu_2 & \begin{bmatrix} \langle u \rangle \delta \end{bmatrix} \\
\mu_3 & \begin{bmatrix} \langle u \rangle C \end{bmatrix} \\
\end{align*}$$

This is motivated by the fact that both scalar and non-scalar complements featured in polarity sensitive and scalar expressions. The novel possibility of non-scalar hosts associating with $\mu$ requires us to posit the relevant un-interpretable $[u\delta]$ feature on the $\mu$ particle, by virtue of which the negative polarity system arises automatically, as per the predictions of Chierchia’s (2013) system. Since the union of both the $\delta$ and $\sigma$ alternatives constitute
the entire contextual domain, $C$, the diachronic change can also be stated $(\mu_2)$ in terms of the generalisation of the $\sigma$-feature in OJ to the $C$-feature in CJ. A stronger view of the the loss of restriction on associating structural contexts of $\mu$ in CJ, and in fact modern Japanese, would be $(\mu_3)$ to assume complete loss on the kinds of complements $\mu$ was associate with. The features restriction, as discussed, refers to the scalar/non-scalar property of the $\mu$-associate. The $\mu$ particle is assumed to be restricted to $\varepsilon$-type associates (as discussed in Mitrović and Sauerland 2016).

Therefore, the relevant change is that in the type of inference that $wh$-$\mu$ expressions carried. In OJ, such expressions are analysed as (positive) SIs, while the Classical Japanese $wh$-$\mu$ expressions were, or at least could be, NPIs under negation.

In (34a) and (34b) I present this view of change in inferential procedure due to featural change, which I analyse as the signature property of the grammaticalisation of $\mu$ in Japonic.

(34) a. OJ: $\neg \rightarrow \forall \vdash \forall (SI)$

\[
\begin{align*}
\text{i. } & \text{EXH}_{\muP} \vdash_{[\muP : \sigma]} \neg \muP & \rightarrow \rightarrow \text{SI}^+ \\
\text{ii. } & \text{EXH}_{\sigmaP} \vdash \neg \ldots \lfloor \exists \sigma \lor \muP \rfloor
\end{align*}
\]

b. CJ: $\forall \rightarrow \neg \vdash \neg \exists (\text{NPI})$

Note that in Chierchia’s (2013) system, which is assumed here, both (positive) SIs and NPIs are derived using the same apparatus.
I address later the technical details of deriving the meanings above. In the next section, I proceed to apply this technical analysis to the IE superparticle system by showing how this application derives the desired natural classes of NPIs and universal (PPI) terms.

Fig. 1 schematises the diachrony of μ in Japonic along with some inferential connections between the meanings that arise in its history.
3.2.2 The Rise of Polarity Sensitivity in Indo-European

The argument for the rise of polarity sensitivity, and the subsequently single-layer of exhaustification over a DE context, is not made as straightforwardly as for Japonic, simply because the IE language family is less homogenous.

The μ has existential quantificational force and forms NPIs only in the Satem group of IE languages, such as Slavonic and Indo-Iranian. In the Celtic, Italic, and Anatolian language families, however, μ forms universal quantifier expressions (as well as FCIs, which I address elsewhere as they outside the scope of this section). To reiterate some core contrastive data, exemplifying this quantificational split, compare the meaning of a wh-μ expression in Sanskrit (35) and Gothic (36), where one is a polarity sensitive expression and another a (polarity ‘insensitive’) universal expression.

Given the phylogenetic evidence for the Anatolian branch having split off from the Indo-European ‘core’ the earliest, it is reasonable to assume that Anatolian, therefore, reflects a more retentive character of μ. The fact that both Indo-Iranian and Slavonic (among other Satem language families) cannot express universal quantificational expressions using a conjunctive
\(\mu\)-particle, while the inverse is true for Centum languages such as Hittite, Gothic, and Latin, leads one to conclude that the primary meaning of \(\mu\)-based \(wh\)-expressions was universal quantificational, rather than than of a negative polar existential.

This change from universal quantificational expressions to existential polar ones is, in the system I am advocating, modelled as loss of iterative exhaustification and its replacement with a non-iterative exhaustification restricted to DE environments.

In the next section, I turn to formalising some diachronic trends and principles of change that have been detected in the history of Japonic and Indo-European.

3.3 Formalising Changes

The primary meaning of \(\mu\) in Japonic, analogous to \textit{even}, has—as shown in previous subsections—descalarised and gave rise non-scalar meanings, such as universal quantification, NPIs, and additives. In the Chierchia’s (2013) system I assume here, there is a way of capturing this diachronic change as a change of mode of exhaustification, from \textit{even}-type to the \textit{only}-type. The other view, and the one I subscribe to and advocate here, is more economical and finds the locus of change in the structural representation of LF, while collapsing one mode of exhaustification into the other. Let me first discuss the first, and most obvious, theory of change.
3.3.1 First Approach: The Diachronic Change in the Mode of Exhaustification

On a more general, and theoretical, level, this section concerns the question whether *even*-type inferences can be derived using the same technology and exhaustification apparatus as polarity-sensitive phenomena. In Chierchia’s (2013), *even*-meanings obtain through an application of an entirely different type of exhaustification operator: $E(\text{even})$-type $\text{exh}$, as opposed to the $O(\text{only})$-type $\text{exh}$. (Note that ‘$<_\pi$’ is a contextually relevant probability measure.)

(37) Chierchia’s (2013) two types of exhaustification operators:

\begin{align*}
a. & \quad O_\Delta(p) = p \land \forall q \in \Delta[p \nRightarrow q] \\
b. & \quad E_\Delta(p) = p \land \forall q \in \Delta[p <_\pi q]
\end{align*}

As the primary aim of this paper is to provide a theory of diachronic semantics which would shed light onto the relation between positive universal, polar-sensitive negative indefinites (NPIs), additives, and *even*-type meanings (OJ), I am led to assume that variation and change in the logical form of these expressions is systematic.

In Chierchia’s (2013) system, the diachronic semantics of the Japonic $\mu$, pertaining to the semantic change ($\gg$) from Old to Classical Japanese, has the locus in the type of exhaustification brought about by the $\mu$ particle, as per (38).

(38) $\left[ c_p E \cdots [\mu [\text{XP} \!]] \right] \gg \left[ c_p O \cdots [\mu [\text{XP} \!]] \right]$

It is unclear that, if any, natural principle of change would underlie the shift from $E$-type to $O$-type exhaustification that is evident in Japonic. The
alternative trigger, $\mu$ or a PSI like any or irgend-., does not select the mode of exhaustification in the system as it stands. It is, therefore, unclear for these reasons how a diachronic mechanism, such as the one in (38) could be systematically accounted for in a way which would yield any predictive power.

Even on a more theoretical level, and not just at the level of diachronic finding and deriving diachronic explananda, I turn in 3.3.2 to dissecting even, being an overt version of the covert $E$-exhaustifier, and breaking it down into a recursive $O$-type exhaustifier coupled with a total-ordering map. This has the potential of economising the inventory of exhaustification procedures as the difference between $E$- and $O$-mode of exhaustification may be eliminated.

For these reasons, I now turn to exploring the other approach.

3.3.2 Second approach: dissecting even

If a meaning can be dissected into compositional components that in concert return the relevant meaning, then structural dissection can be relegated to diachronic morpho-syntax.

This section explores one such avenue for the meaning of even, to which I relate the primary meaning of $\mu$ in Japonic, as demonstrated in §3.1. In the following, I use lattice-theoretic apparatus ($\prec$ is a covering relation).

After additive meaning obtains, via a recursive application of $\text{exh}$ (or a single application, assuming $\text{exh}$ is defined in this case over a set of pre-exhaustified alternatives), the partially ordered alternative set is mapped onto a totally ordered set (chain). See Stanley (1984) or Wild (2005), among others, for details, which are not relevant for the basics of the proposed mechanism.
(39) \[ [[\text{even}]] \]

\[
\begin{array}{c}
\neg \text{EXH}(p) \circ \Omega \\
\neg \text{EXH}_{\delta/\sigma}(p) \\
\text{EXH}_{\sigma/\delta} \\
\text{EXH}_{\delta/\sigma}(p) \\
\mu P \\
\mu_{\delta/\sigma} P
\end{array}
\]

a. \( \text{EXH}_C(p) = p \land \forall q \in \bigvee (\mathcal{A}_C(p))[\neg q] \), where \( C \) ranges over a semi-
lattice of \( \sigma \)- and \( \delta \)-alternatives.

b. \( \text{EXH}^R(p) = \text{EXH}(\text{EXH}(p))(p) = p \land \neg \text{EXH}(p) \)

c. \( \Omega = [C] \)

d. \( r : \mathfrak{A} \mapsto C \mid C = \{0, \ldots, n\}, \text{s.t.,} \)

i. Let \( r \) be an order preserving bijection from the lattice \( \mathfrak{A} \) onto

rank-selected (maximal) chain \( C \) of \( \mathfrak{A} \):

\[
\begin{array}{c}
\triangleright r(x) = \begin{cases} 0 & \text{for all minimal } x, \text{ i.e. iff } \exists y[y < x] \\
      n & \text{for all maximal } x, \text{ i.e. iff } \exists x[y < x]
\end{cases}
\end{array}
\]

\[
\triangleright r(y) + 1 \iff \exists y[y < x]
\]

e. floor-function: \([p]\), takes an element \( x \) of the chain \( C \) and returns

the greatest element \( y \), s.t. \( y \leq x \). (Negation reverses the func-
tion, i.e. its dual (ceiling-function) obtains.)

f. \( h : \Omega^{-1}(0) \) defined on a two-element chain \( C_2 = \{\{0, 1\}, \leq\} \)

Note that the LF dissection in (39) is not a hardwired lexical entry for OJ \( \mu \),
but rather the representation of the components that OJ \( \mu \) brings into play.

The evidence for the proposed dissection of \( \text{even} \) or \( \text{even} \)-like operators, which
relies on its additive component being derived from recursive application
of the \( \text{exh} \)-operator, is also supported by empirical motivation from English
Figure 2: Maps and partial homomorphisms from completely lattice-structured alternatives to a ranked chain and two-element chain.

(Francis 2017, and those cited) and Hungarian (Szabolcsi, 2017b). In English, the additivity of even is more or less standardly assumed and taken for granted (cf. Rullmann 1997, int. al. for an opposing discussion). In fact, obligatory additivity of even derives for free by virtue of even’s bringing into play likelihood comparanda, or at least one such comparandum which is more likely that the assertion.

In fact, the arguments against the additivity of even, which is part of the dissected meaning in (39), may be accounted for by appealing to the syntactic distribution of even and its associates. In this direction, Wagner (2015) defends a generalisation pertaining to the category of even-associates:

(40) Wagner (2015) generalisation:

NP-even is additive, VP-even is not.

Since this paper is primarily focussed on the NP-combining μ particles (see Mitrovic and Sauerland 2016 for background), the additivity of even in nominal structural contexts is thus well established.

Szabolcsi (2017b) introduces novel evidence from Hungarian which are on a par with Ser-Bo-Croatian with regards to the fact that two particles must act in tandem to deliver the even-meaning of low-rankedness (or related con-
cessive meaning, cnccs) or unlikelihood.

(41) ‘Even Mary’

a. Hungarian

i. még Mari is
   **EVEN** **ADD**

ii. még/akár csak Mari is
   **EVEN**<sub>1</sub> **ONLY** **ADD**

b. Ser-Bo-Croatian

ˇcak/makar i Marija
   **EVEN**/**EVEN**<sub>cnccs</sub> **ADD**

Thus, in these NP-associating contexts, **even** is encoded by two particles in Hungarian and Ser-Bo-Croatian, in Hungarian up to three particles (41a-ii) are lexicalised. The Ser-Bo-Croatian **i** and Hungarian **is** are independently additive. The **még** and **akár** are, according to Abrusán (2007) even-style particles (which is what Szabolcsi 2017a also assumes), hence labeled **even**.

Note that the Hungarian **csak** and Ser-Bo-Croatian **ˇcak** are cognate borrowings from Turkish (Skok, 1971: 289). The original borrowed form, **ˇcak**, had (and still has in modern Turkish) the low-likelihood meaning **even**. It is curious that Hungarian borrowed, or developed, the **only**-type meaning. In this connection, the **only**-type particle **sam**- ‘only, alone’ shows interesting connection in Ser-Bo-Croatian also, bringing the Hungarian and Slavonic meanings of **csak/ˇcak** (only/even, resp.) synchronically closer in their distribution (42). Also note, however, that the particles may also co-occur (43) which is in line with our proposed dissection:

(42) Skok (1971: 289)

a. **ˇcak** do zore
even to dawn. **ACC**
   ‘**even** till the dawn’

b. do zore **same**
to dawn. **ACC** only/alone. **F**
   ‘**even** till the dawn’
Frurthermore, Ser-Bo-Croatian makar is cognate, and (as far as I am able to establish) synchronically rquivalent to the Slovenian concessive scalar particle magar(i) (Crnič, 2011).

Skok (1972: 359) notes that makar is “of unknown origin from the domain of syntax,” as it’s found in Romanian măcar, and even modern Greek μακάρι. One analysis that Skok entertains is the coupling of the neo-Persian negative morpheme ma and agar (from Old Persian *hakaram, ‘once’; cf. Turkish meğer). Without reference to the original direction or form of the borrowing, I conjecture that makar/magar(i) are cognate with Hungarian particle sequence még and akár. Historically, the Turkish contact is relevant and likely. This is parallel to čak/csak and its potential borrowing from Turkish (cf. čak and the discussion above).

3.4 Parametrising the semantic incarnations of μ

In this final subsection, I attempt to integrate the previous discussion into a coherent explanatory model of diachronic variation and change of μ, deriving from the proposed dissection of even-type meaning, which I reproduce below.

(44) Parametric conditions on s scalar-additive template for μ:
The idea behind this approach is based on the assumption that a carefully dissected meaning of Japonic μ provides sufficient means of deriving not only the non-scalar Indo-European character of μ but also the diachronic procedures that yielded the semantic shapes of μ in non-primitive language states.

The first parameter set, P₁, is featural and, on the one hand, distinguishes between strictly scalar and non-scalar association (μ-complementation) as given in (45) for P₁a.

(45) \( P₁a : μ[±σ] = \begin{cases} + & \text{Old Japanese} \\ - & \text{rest (both scalar and non-scalar.)} \end{cases} \)

On the other hand, parameter P₁b distinguishes between (indeterminate, \text{indet}) wh-complements and (determinate) DP complements (45). This differentiates the Old Japanese type of μ-encoded meanings from the rest, in
both comparative and historical terms.

\[(46) \quad \mathbf{P}_{1b} : \mu[\pm \text{INDET}] = \begin{cases} + \text{ indeterminate quantificational meanings} \\ - \text{ additive meanings} \end{cases} \]

The second parameter, \( \mathbf{P}_2 \), is also related to morpho-syntax and distinguishes between purely \( e \)-type \( \mu \) particles in the nominal domain and the \( \mu \) particles in the clausal domain \((47)\). One such example would be Gothic and Modern Greek, as discussed in Mitrović \((2018b)\). Evidence from Gothic, and even Old Irish, by virtue of lacking an exclusively DP-conjoining character, suggest a fixed position for the \( \mu \) particle. This position is independently shown to be a clausal Focus head, which supports the view that the IE \( \mu \) superparticle was reanalysed as an element in the C-system (see Mitrović \((2018b)\) for details and an account for this.)

\[(47) \quad \mathbf{P}_2 : \mu[\pm \text{D-LEVEL}] = \begin{cases} - \mu \text{ in the C-system: Gothic, Modern Greek} \\ + \mu \text{ in the C-system: rest} \end{cases} \]

\( \mathbf{P}_3 \) in \((48)\) regulates Proper Strengthening \((10)\) as a presuppositional condition of \( \mu \)-expressions. At this level, I am no longer distinguishing between languages but meanings that are allowed if this parameter is (not) set. The meaning that \( \mathbf{P}_3 \), based on PS, filters is that of (non-) polarity-sensitivity. Therefore, if PS operates, and the relevant \( \mathbf{P}_3 \) is set positively, ‘+’, then \( \mu \) will not, or does not, partake in forming NPIs.

\[(48) \quad \mathbf{P}_3 = [\pm \text{PS}] = \begin{cases} + \text{ no polarity-sensitive meanings} \\ - \text{ polarity-sensitive meanings allowed} \end{cases} \]

\( \mathbf{P}_3 \) can be seen as an economy condition: if all meanings derived by \( \text{exh} \) are presupposed to be informationally stronger than the (input) assertion,
then the only means available for strengthening is the reapplication of exh, which leads me to stating $P_4$ in (49).

$P_4$ is iterativity allowed or obligated?

\[(49) \quad P_4 = [\pm \text{iter}] = \begin{cases} - & \text{non-strengthened meanings only} \\ + & \text{additivity, free-choice, universal quantification possible} \end{cases} \]

Similarly to $P_3$ in (48), $P_4$ will allow additive, universal quantificational and free-choice meanings, to the exclusion of NPIs. If additivity can be separated from other purely quantificational $\mu$-expression by the $[-\text{indet}]$ parameter setting, then how does our analysis discriminate between vanilla universal quantifiers and the universal FCIs? In this regards, we follow Dayal (2009) and adopt her fluctuation constraint (13), as discussed earlier.

Therefore, $P_5$ can be stated in terms of Is fluctuation obligated?

\[(50) \quad P_5 = [\pm \text{fluc}] = \begin{cases} - & \text{universal quantification} \\ + & \text{free-choice} \end{cases} \]

Also note that $P_2$ and $P_5$ are connected, although the former is a purely syntactic and the latter a purely semantic parameter. The grammaticalisation of a FC-marker in the C-system guarantees ($P_2 : -$) Fluctuation and the wide-scope of $\mu P$.

The final parameter I state in (44) is $P_6$ in (51) has to do with the non/scalar output and is related to the first parameter on non/scalar complementation. It can be informally read as ‘Is scalar meaning obligated?’

\[(51) \quad P_6 = [\pm \text{scal}] = \begin{cases} + & \forall x, y \in \mathcal{A}(\mu P)(x < y \lor y < x) \\ - & \text{else, non-scalar meaning} \end{cases} \]

Note that the $\Omega$ and $h$ functors are motivated empirically on the basis of evidence for $even$ associating in questions, as discussed in Iatridou and Tat-
Parameter $P_6$ will thus allow for scalar and even-type meanings of $\mu$-expressions, which captures not only the character of Old Japanese but also the seeming universality of the scalar ambivalence of additives.

(52) A parametric hierarchy for the meanings of $\mu$:

The $\Omega$ and $h$ differ, basically, only with respect to whether they only rank the alternatives via a map to a chain or whether they homomorphisms onto a two-membered chain of un/likelihood. See Fig. 2
4 CONCLUSIONS & OUTLOOK

I have attempted to sketch a diachronic theory of \( \mu \)-meanings: how they are born and how they shift. Two such shifts were demonstrated. In Japonic, the meaning of \( \mu \) is that of scalar particle: the best means of capturing this distribution is by positing, in Minimalist terms (Chomsky, 1995) an uninterpretable scalar (\( \sigma \)) feature on the Old Japanese \( \mu \). Once this restriction to scalar complements is lost in Classical Japanese, non-scalar meanings arise in Japanese. One such meaning is that of neagative polarity or polarity sensitivity more generally. In order to obtain this meaning, I proposed, in line with Chierchia’s (2013) system, that the relevant feature allowing access to non-scalar sub-domain (\( \delta \)) alternatives is innovatively born on the Classical Japanese \( \mu \) head. This polar meaning is also secondary in Indo-European. I supported this view with independent phylogenetic arguments according to which the two groups, the one in which \( \mu \) encodes NPIs and another in which \( \mu \) encodes universal quantifiers, are differently retentive with regard to the original (Proto) Indo-European character of \( \mu \).

In the last section, I sketched a view of the causes of diachronic shifts that are detectable in Japonic and Indo-European. These causes can be tangentially gathered under the umbrella of economy conditions (explored in greater detail elsewhere). One such change in adhering to economy was is synchronically by the grammatical use of the \text{exh} operator in ‘last resort’ derivations/interpretations: if an exhaustified proposition does not make sense, then run exhaustification once more. I stated this as the iterativity parameter \( P_4 \). Another example of an economical principle at play was tied to the presuppositional condition that the application of \text{exh} has to
Mitrović

strengthen the meaning in which it plays a role. If this condition, stated as parameter $P_3$, is part of the grammar, then $\mu$-encoded NPIs are predicted to not exist in a given language.

The latter parameter also leads us to an implicational universal which is seemingly trivial: if an expression obtains in non-fluctuational contexts, then it is also admissible in fluctuational contexts. I have not paid particular concern to unviersal FCIs in this paper so I leave the details of this universal, and the precise relation of Fluctuation to the overall principles of $\mu$, for future work.

What remains to be worked out, on a more general level, is the connection between the proposed parameters as well as their wider empirical examination. It appears to me that an explanatory relegation of historical phenomena to principles of economy and efficient computation (both derivational and interpretational) are a desirable consequence of a diachronic semantic theory.

REFERENCES


