

Appendix A. Fundamentals of the Emu query language

The purpose of this Appendix is to set out in some further detail the different kinds of queries in the Emu query language.

A.0 General

A query to the Emu query language (Emu-QL) consists of one or more arguments. The syntax for an argument is $T = a$ where T is a tier and a an annotation at that tier. The meaning of $T = a$ is: find the annotations a at tier T ¹. In the examples presented below, the queries are made with respect to the path shown in Fig. A.1 of the downloadable database `andos1` (a fragment of the Australian National Database of Spoken Language - see Millar et al, 1994).

In the following description of the various commands, the square bracket and forward-slash notation will be used to denote annotations at the `Phonetic` and `Phoneme` tiers respectively e.g. `[m]` and `/m/` for "m" annotations at these tiers. In addition, a *strong syllable* will be used to denote an `S` annotation at the `Syllable` tier, an *accented word* an `S` annotation at the `Accent` tier, an *L-phrase* denotes an `L-` annotation at the `Intermediate` tier.

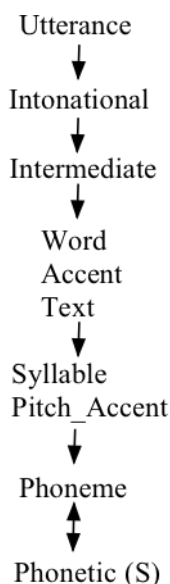


Fig. A.1. One of the paths in the downloadable `andos1` database from which queries are made in these examples. `Phonetic` is a segment tier from which other tiers inherit their times. `Word`, `Accent`, `Text` are linearly linked as are `Syllable` and `Pitch-Accent`. The relationship between a parent and child tier is hierarchical in all cases. With the exception of `Phoneme:Phonetic`, it is also one-to-many.

A.1 Simple queries

Note: See 4.2 of Chapter 4 for accessing features using the `&` and `!=` operators

`[m]`
`Phonetic = m`

`[m] or [n]`
`Phonetic = m | n`

All annotations except `[m]` or `[n]`

¹ When a consists of numbers or special characters, the annotation needs to be quoted (e.g. $T = '1'$). It is not possible to search for apostrophes.

Phonetic != m | n

All syllables (assuming none have been annotated with x)

Syllable != x

A.2 Sequence queries

Syntax: [T = a -> T = b]

Meaning: A sequence of annotations a and b both from tier T

Conditions: Sequence queries can only be made from the same tier or from linearly linked tiers.

[m I]

[Phonetic = m -> Phonetic = I]

[m] before [I]

[#Phonetic = m -> Phonetic = I]

[I] following [m]

[Phonetic = m -> #Phonetic = I]

[m I n]

[[Phonetic = m -> Phonetic = I] -> Phonetic = n]

John could lend

[[Text = john -> Text = could] -> Text = lend]

the followed by any two words followed by managed

[[[Text = the -> Text != x] -> Text != x] -> Text = managed]

A.3. Queries from tiers that stand in a linear relationship to each other

Syntax: T = a & U = b

The orthography of all function words

Text != x & Word = F

Accented content words i.e. any annotation at the Text tier linked to a C annotation at the Word tier and to an S annotation at the Accent tier.

Text != x & Word = C & Accent = S

A.4. Queries from tiers that stand in a non-linear relationship to each other

/p/ in strong syllables (any p annotation at the Phoneme tier linked to an S annotation at the Syllable tier)

[Phoneme = p ^ Syllable = S]

Any syllable containing /p/

[Syllable != x ^ Phoneme = p]

Equivalent to the above

[Phoneme = p ^ #Syllable != x]

The phrase tone(s) associated with H* pitch-accented *price* or *space* (find annotations at the Intermediate tier that include either *price* or *space* at the Text tier that are linked to H* at the Pitch_Accent tier)

$[[\text{Intermediate} \neq x \wedge \text{Text} = \text{price} \mid \text{space}] \wedge \text{Pitch_Accent} = \text{H}^*]$

A.5. Position

Syntax: $\text{Start}(T, U)=1$

Meaning: annotations at tier U that are at the beginning of tier T

Condition: T dominates U

Word-initial syllable

$\text{Start}(\text{Word}, \text{Syllable})=1$

Phrase-final syllables: i.e., annotations at the Syllable tier that are final with respect to the Intermediate tier

$\text{End}(\text{Intermediate}, \text{Syllable})=1$

Word-medial or word-final syllables

$\text{Start}(\text{Word}, \text{Syllable})=0$

Word-medial syllables

$\text{Medial}(\text{Word}, \text{Syllable})=1$

A.6. Position and linear links

Syntax: $T = a \ \& \ \text{Start}(U, V)=1$

Meaning: Find annotations at tier V which occur at the beginning of tier U and which are linearly linked to a annotations at tier T

Condition: U dominates V. Also T and V are of the same tier or two linearly linked tiers.

Note: $\text{Medial}(U, V)$ for medial, $\text{End}(U, V)$ for end.

Syllable-initial /p/ phonemes

$\text{Phoneme} = p \ \& \ \text{Start}(\text{Syllable}, \text{Phoneme})=1$

Word-final /m/ phonemes

$\text{Phoneme} = m \ \& \ \text{End}(\text{Word}, \text{Phoneme})=1$

Word-medial or word-final strong H* syllables

$\text{Syllable} = S \ \& \ \text{Pitch_Accent} = \text{H}^* \ \& \ \text{Start}(\text{Word}, \text{Syllable})=0$

A.7. Position and non-linear links

Syntax: $[T = a \wedge \text{Start}(U, V)=1]$

Meaning: Find annotations at tier V which occur at the beginning of tier U and which are non-linearly linked to a annotations at tier T

Condition: U dominates V. Also T and V are non-linearly linked.

/p/ phonemes in word-initial syllables

$[\text{Phoneme} = p \wedge \text{Start}(\text{Word}, \text{Syllable})=1]$

Phonemes that are not in word-final syllables

[Phoneme != x ^ End(Word, Syllable)=0]

A.8. Number

Syntax: Num(T, U)=4

Meaning: Find annotations at level T linked to a sequence of 4 annotations at level U

Condition: T and U are non-linearly linked

Note:

>	more than
<	fewer than
<=	fewer than or equal to
>=	more than or equal to
=	equal to
!=	not equal to

Trisyllabic words

Num(Text, Syllable)=3

Syllables with 6 or more phonemes

Num(Syllable, Phoneme)>=6

A.9. Number and linear links

Syntax: T = a & Num(U, V)=4

Meaning: Find the annotation at tier T linearly linked to tier U which includes 4 annotations at level V

Condition: T and U are linearly linked

The words (annotations at the Text tier) consisting of more than or equal to 5 phonemes

Text != x & Num(Text, Phoneme)>=5

H* syllables consisting of more than 3 phonemes

Pitch_Accent = H* & Num(Pitch_Accent, Phoneme)>3

A.10. Number and non-linear links

Syntax: [T = a ^ Num(U, V)=4]

Meaning: Find the annotation at tier T non-linearly linked to tier U which is non-linearly linked to 4 annotations at level V

/m/ phonemes in trisyllabic words

[Phoneme = m ^ Num(Word, Syllable)=3]

Weak syllables in words or three or fewer syllables

[Syllable = W ^ Num(Word, Syllable)<=3]

H- intermediate phrases containing words of 2 or more syllables

[Intermediate = H- ^ Num(Text, Syllable)>2]

A.11. Combination queries (non-linear and sequence)

Annotation a at tier T precedes annotation b at the same tier and a is non-linearly linked to c at tier U.

[[#T = a -> T = b] ^ U = c]

Annotation b at tier T following a at the same tier and b is non-linearly linked to c at tier U.

[T = a -> [#T = b ^ U = c]]

/m/ in a sequence /m I/ in a strong syllable

[[#Phoneme = m -> Phoneme = I] ^ Syllable = S]

/I/ in a sequence /m I/ in a strong syllable

[Phoneme = m -> #Phoneme = I] ^ Syllable = S]

Weak syllables containing /p/ such that /p/ follows /s/

[Phoneme = s -> [Phoneme = p ^ #Syllable = W]]

Strong syllables containing /p/ phoneme such that these strong syllables precede strong syllables

[Syllable = S ^ Phoneme = p] -> Syllable = S]

As above, but return the /p/ phonemes (i.e., /p/ phonemes in strong syllables such that the strong syllable precedes a strong syllable)

[Syllable = S ^ #Phoneme = p] -> Syllable = S]

A.12. Combination queries (non-linear and sequence and linear)

Words that beginning with a schwa

[Text! = x ^ Phoneme = @ & Start(Text, Phoneme)=1]

Word-initial /m/ preceding /I/ in a strong syllable

[[#Phoneme = m & Start(Word, Phoneme)=1 -> Phoneme = I] ^ Syllable = S]

As above, but return the corresponding word (the corresponding annotation at the Text tier):

[[[Phoneme = m & Start(Word, Phoneme)=1 -> Phoneme = I] ^ Syllable = S] ^ #Text != x]

The orthography (annotations at the Text tier) of all words of two or more syllables which have a schwa in the first syllable and which precede the

[[#Text != x & Num(Text, Syllable)>=2 ^ [Phoneme = @ ^ Start(Text, Syllable)=1]] -> Text = the]

Words including a syllable-initial and intermediate phrase-final /f/ phoneme in L-intermediate phrases of more than 3 syllables

[Text != x & End(Intermediate, Text)=1 ^ Intermediate = L- & Num(Intermediate, Word)>3] ^ Phoneme = f & Start(Syllable, Phoneme)=1]