

# Text-setting in Kaytetye

Nay San and Myfany Turpin  
*Stanford University and University of Sydney*

## 1 Introduction

Singing is a universal human activity. Across the vast range of song traditions throughout the world, native speakers have consistent intuitions about how the syllables in a given line of song text should be set to the tunes and/or rhythms within their various song traditions. Remarkably, even across many traditions and languages, there appear to be consistent preferences in text-setting, for example, for how prominent parts of text such as stressed syllables should be aligned to prominent parts of musical rhythm such as downbeats. Similar preferences also exist in poetry.

While the basis for these preferences in text-setting has been the subject of some research in phonology and metrics (e.g. Halle & Lerdahl, 1993; Dell & Halle, 2009; Hayes, 2009), the song traditions analysed have not involved significant changes to the song texts when sung. If presented with such a tradition, then, it is natural to ask the following questions: Why does the metrical template of the song tradition require such significant changes to the sung text? What properties of the language permit such changes? How do the rhythmic requirements and the set of permitted changes interact such that a specific type of change is expected in a given rhythmic context for a given text but not in any others? More broadly, what do such interactions reveal about the relation between language and music?

In this paper, we present an analysis of text-setting in *Akwelye* [a'kɔlə] — a set of ceremonial songs traditionally sung and passed on orally by groups of Kaytetye-speaking women in Central Australia. Australian Aboriginal songs are renowned for the degree to which they diverge from speech (Clunies Ross et al., 1987; Sutton, 1987). *Akwelye* has been documented in detail by Turpin (Turpin & Ross, 2004; Turpin, 2005, 2007a,b). As shown in (1a–c), the sung forms of *Akwelye* song texts (shown in bold) may exhibit certain systematic differences — henceforth ‘non-faithfulness’ — compared to their spoken forms (shown in square brackets) in everyday spoken Kaytetye and neighbouring languages, from which much of the song’s vocabulary is drawn. In particular, we consider the deletion and addition of segments. We do not consider the modification of quality, which is motivated by a different set of constraints.

(1) Attested relations between sung and spoken forms in the *Akwelye* song tradition.

	a. Vowel Deletion	b. Consonant Epenthesis	c. Syllable Transfer	d. Faithful
Sung	<b>'wɪɫɔwə 'rəŋə</b>	<b>'lɪɫpɔɫi 'məɪa</b>	<b>'nəpəɟipəla</b>	<b>'kurɪmpi 'kurəi</b>
Spoken	[a'wɪɫɔwə 'rəŋə]	[a'ɫpɔɫi 'pɪməɪə]	[a'pəɟipələnə]	['kurɪmpə 'kurə]
Underlying	/awɔlə-wə arə+ŋə/	/ɫpɔɫə əɟpɪməɪə/	/apəɟɪpə-lə-nə/	/kwərəɟɪmpə kwərə/
Gloss	‘ceremony-DAT put+RPST’	‘leaf dew’	‘lizard-ERG-VOC <sub>1</sub> ’	‘ancestral.woman girl’

In (1a), the initial vowel [a] in the spoken form is omitted from the sung form. In (1b), an initial consonant [l] not present in the spoken form is epenthised to the beginning of the sung form. In (1c), a transposition affecting both edges of the word is made: the final syllable [nə] in the spoken form is transferred to the line-initial position in the sung form (note that [nə] + [a] forms [na] through standard vowel hiatus resolution in Kaytetye). Yet, not all sung forms in *Akwelye* exhibit non-faithfulness; as shown in (1d), the sung form may be relatively faithful to its spoken form. The non-faithfulness of (1a–c) raises the question: what linguistic and metrical contexts condition non-faithfulness in *Akwelye* text-setting?

We will show that both the presence and type of non-faithfulness found within the sung forms of *Akwelye* are the result of an interaction between the phonological properties of Kaytetye and the rhythmic requirements of the *Akwelye* song set. For example, Kaytetye has fixed word stress which is invariably placed on the first syllable of a word with an onset. All but two of the non-faithful sung lines are associated with song texts that begin with an unstressed utterance-initial vowel, e.g. the initial [a] in (1a) [a'wɪlɔwa 'rəŋə]. Singing this line faithfully would involve alignment of the unstressed vowel [a] to the downbeat (a rhythmically strong location), and hence one of the non-faithful options is preferred.

Additionally, among the non-faithful options, the deletion of the initial unstressed vowel [a] is preferred in the case of (1a) given the morpho-phonological structure of the song text and rhythmic requirements of *Akwelye*. As introduced below, elements within complex words may have one of two morphological relations: 'word-level' or 'root-level', as is the case for many Australian languages (Baker, 2014:157). The transferring of the utterance-final syllable to the initial position [ɪŋawɪlɔwa 'rə] is dispreferred, as the final-syllable [+ŋə] is a root-level suffix. Consequently, the two remaining options for (1a) are vowel deletion or consonant epenthesis. The latter is dispreferred, as the epenthesis of a consonant [l] for example would have resulted in a 6-syllable utterance [l'awɪlɔwa 'rəŋə], composed of a 4-syllable word followed by a 2-syllable word. For reasons explicated further below, such a composition does not fit as optimally to the metrical template of *Akwelye*, which has a strong preference for a 3-syllable word followed by a 2-syllable word — i.e. [wɪlɔwa 'rəŋə], which is achieved by omitting the initial vowel [a]. In this way, vowel deletion, consonant epenthesis, and syllable transfer permit the optimisation of variable-length inputs for the preferred rhythmic setting.

While the metrical template of *Akwelye* is fixed (4 grid positions per measure), song texts may be set to as few as 2 measures or as many as 4 measures, given the variable length of the input. In conjunction with the ability to modify these inputs in various ways, there exist many possible settings for each song text. For our analysis, we exhaustively generated all possible settings using a method based on integer compositions: all the ways an integer  $n$  may be expressed as a sum of smaller positive integers, e.g. compositions of 6, or  $C(6) = \{4+2, 4+1+1, \dots, 1+1+1+1+1\}$ . Given a song text of  $n$  syllables and a metrical template with  $m$  positions within each measure, all possible settings of this song text to can be computationally generated using  $C(n)$  and  $C(m)$ . We provide a step-by-step walkthrough of this generation process further below.

The constraint set used in our analysis is able to account for 71 of the 89 *Akwelye* song lines. Several were existing constraints which were either adapted from previous research (e.g. SALIENT from Kiparsky, 2006) or adopted as-is (e.g. \*STRESSLESS IN STRONG from Hayes, 2009). Additional constraints are motivated on the basis of language-specific properties of Kaytetye, or meter-specific properties of *Akwelye*. In other words, we show that the seemingly idiosyncratic nature of text-setting strategies in *Akwelye* can be accounted for through a relatively generic set of constraints (even when thousands of competing candidates are considered), reflecting many of the same fundamental processes that govern the interaction of language, meter, and music.

## 2 Kaytetye prosody and morphology

Kaytetye is analysed as having only two vowel phonemes /a/ and /ə/ (Koch, 1984), and an unusually large 52-phoneme consonantal inventory. Surface vowel qualities such as [ɪ], as within the name of the language: *Kaytetye* /kajtɛcə/ ['kajtɛcə], arise from allophonic variation (e.g. /ə/ is fronted to [ɪ] when preceding a palatal such as /c/). Vowels are contrastive in initial and medial positions, but not in the final position. Standard analyses of Kaytetye regard all words as ending in a non-contrastive final vowel /ə/ (Koch, 1984, 2004; Turpin & Ross, 2012). Initial and/or final vowels can be variably omitted in colloquial speech, e.g. /ələkə/ 'dog' [a'ləkə], ['ləkə], [a'lək], ['lək]. At word and morpheme boundaries, the final vowel of the preceding element is elided through vowel hiatus resolution when preceding a vowel-initial morpheme, e.g. /ələkə/ 'dog' + /-akakə/ 'PROP(RIETIVE)' → [a'ləkəkəkə] 'having dogs', i.e. \*[a'ləkəəkəkə].

Kaytetye has fixed stress, like many Australian languages. Unlike most Australian languages, however, Kaytetye has word-initial vowels. Stress assignment in Kaytetye is onset-sensitive and footing proceeds trochaically from left to right starting with the first syllable with an onset, e.g. a('lɛ.kə) 'dog' vs. ('kɪ.rə) 'girl'; word final vowels can not be stressed, e.g. \*['ku.rɪm'bə]. Thus, in Kaytetye, the word-initial onsetless syllable can never be stressed, and we will see below how stress is relevant to text-setting.

In Kaytetye, elements within complex words may have one of two morphological relations: 'word-level' or 'root-level', as is the case for many Australian languages (Baker, 2014:157). The motivations for this distinction are complex and beyond the scope of this introduction (see Baker & Harvey, 2003; Baker, 2014),

though given their usage in the Australianist literature, we follow this terminology in this paper, denoting word-level morpheme boundaries with ‘-’ and root-level morpheme boundaries with ‘+’. For the analysis of *Akwelye*, it suffices to note that word-level suffixes such as the locative [-lə] only attach to stems which may form independent words, e.g. [a'ləkə-lə] ‘dog-LOC’ vs. [a'ləkə] ‘dog’. Root-level suffixes, on the other hand, predominantly occur with bound verb stems, e.g. [a'larə+nkə] ‘hit+PRESENT’ vs. [a'larə+ŋə] ‘hit+RECENT.PAST’, but crucially \*[a'larə]. In *Akwelye*, syllable transfer (moving a line-final syllable to the line-initial position) is only permissible when the residual stem can form an independent word on its own.

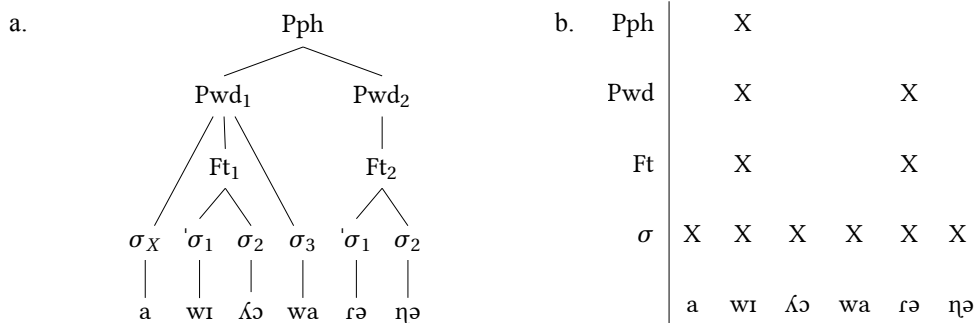
### 3 Exhaustive candidate generation for text-setting

**3.1 Prominence in language and music** Summarising from Kiparsky (2020:§1), we may observe that rhythm in both language and music are formed through alternating prominent and non-prominent beats, e.g. stressed vs. unstressed syllables, and strong beats vs. weak beats. These beats are organised recursively into various rhythmic constituents, each part of some rhythmic domain with a unique rhythmic peak. For example, in prosodic structure, syllables are organised into feet, which are again organised into prosodic words, and so on.

Following Liberman & Prince (1977), we can represent prosodic prominence in a grid denoting a sequence of beats of alternating prominence. For example, shown in (2a) is a tree-based representation of the prosodic structure of an *Akwelye* song text [a'wɪ.ɫɔ.wa 'rə.ŋə] ‘painting on ceremonial body designs’. The same information may also be represented in a grid, as shown in (2b), where the height of the grid of Xs above a syllable denote its relative prominence. Notice that the unstressed syllables [a], [ɫɔ], [wa], and [ŋə] are all only assigned one degree of prominence at the level of the syllable. For the stressed syllables [wɪ] and [rə], however, they receive additional degrees of prominence based on their location within their respective prosodic words and the larger phonological phrase.

Meter in music may also be represented in a grid as a sequence of beats of alternating prominence. Like the prosodic word in phonology, a measure or bar is an organisational unit in music which groups rhythmic constituents within it. Measures have a specified number of beats within it and, shown below in (3), are two measures with two beats within each. As with prosodic structures in language — where we have seen that the prosodic word is a level above the foot and the syllable one below —, beats may be further subdivided into organisational units of smaller sizes or grouped into units of larger sizes. What is most relevant for the

(2) Prosodic prominence of a Kaytetye word in (a) tree- and (b) grid-based representations.



(3) Grid-based representation of musical prominence.

MEASURE	1				2			
	X				X			
Beat	X	X			X	X		
	X	X	X	X	X	X	X	X
Prominence	S	W	M	W	S	W	M	W

the analysis and discussion of text-setting is that each place in the grid is associated with different degrees of prominence: strong, medium, or weak — denoted, respectively, as S, M, and W in (3). Text-setting, then, is a task of finding an optimal alignment between the prominent beats within a metrical template and prominent syllables within the prosodic structure of a song text, subject to the satisfaction of various requirements of the song tradition and the language’s phonology.

**3.2 Integer compositions for exhaustive candidate generation** A composition of an integer  $n$  is a way of expressing  $n$  as the sum of a sequence of integers less than or equal to  $n$  (but greater than 0). For example,  $2+4$  is a composition of 6, as is  $2+2+2$ . The ‘integer compositions of  $n$ ’, written as  $C(n)$ , is the set of all possible ways of expressing  $n$  as such a sum. For 6, there are 32 compositions, thus  $C(6) = \{6, 5+1, 4+2, \dots, 1+1+1+1+1+1\}$ . Note that integer compositions are ordered — thus,  $4+2$  is a different composition to  $2+4$ . Note also that integer compositions may be of different lengths — thus,  $4+2$  is a 2-length composition of 6, while  $2+2+2$  is a 3-length composition.

For text-setting, all possible settings for a sequence of syllables may be exhaustively generated using the integer compositions of two parameters:  $n$  — the number of syllables to be set —, and  $m$  the number of available grid positions per measure. For the *Akwelye* data,  $n$  varies according to the number of syllables in the input song text, e.g.  $n = 6$  for [a'wɪ.ɔ̃.wa 'rə.ŋə] shown in (4a). However, as the metrical template for *Akwelye* has 4 grid positions per measure,  $m$  will invariably be 4, i.e. as shown in (4b). Thus for a given *Akwelye* song text, various candidate settings can be generated using  $C(n)$  and  $C(4)$ , four of which are shown in (5a–d).

- (4) An *Akwelye* song text (a) and the metrical grid to which it is to be set (b).

a. Number of syllables,  $n = 6$

Pph	X					
Pwd	X			X		
Ft	X			X		
$\sigma$	X	X	X	X	X	X
	a	wɪ	ɔ̃	wa	rə	ŋə

b. Positions per measure,  $m = 4$

MEASURE	1			
Beat	X	X	X	X
Prominence	S	W	M	W

- (5) Different ways of setting a song text to the same metrical grid.

a'wɪ.ɔ̃.wa 'rə.ŋə ( $n = 6$ )

$6 = 4 + 2$

a.  $m_1 = 4 = 1 + 1 + 1 + 1$        $m_2 = 4 = 1 + 3$

a	wɪ	ɔ̃	wa	rə	ŋə
---	----	----	----	----	----

b.  $m_1 = 4 = 1 + 1 + 1 + 1$        $m_2 = 4 = 3 + 1$

a	wɪ	ɔ̃	wa	rə	ŋə
---	----	----	----	----	----

c.  $m_1 = 4 = 1 + 1 + 1 + 1$        $m_2 = 4 = 2 + 2$

a	wɪ	ɔ̃	wa	rə	ŋə
---	----	----	----	----	----

d.  $m_1 = 4 = 1 + 3$        $m_2 = 4 = 1 + 3$        $m_3 = 4 = 1 + 3$

a	wɪ	ɔ̃	wa	rə	ŋə
---	----	----	----	----	----

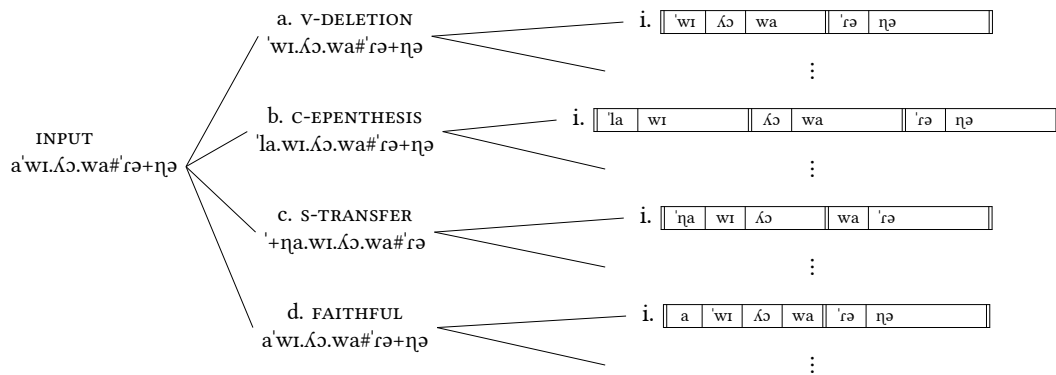
As one concrete example, consider the 6-syllable input text [a'wɪ.ɔ̃.wa 'rə.ŋə], as shown in (5). One possible setting for this sequence of syllables is that shown in (5a), where the first 4 syllables are distributed across the first measure and the remaining 2 syllables are similarly distributed across the second measure. As there are 4 grid positions per measure, there is only 1 possible way in which the first 4 syllables may be set to the metrical template. For the remaining 2 syllables in the second measure, there are 3 ways in which they may be set. One way is to set [rə] to one grid position and [ŋə] to three grid positions, as shown in (5a). Conversely, [rə] may be set to three positions and [ŋə] only one, as shown in (5b). Finally, the two syllables may each be set to two grid positions, as shown in (5c). Thus, 3 possible settings are computed

using a composition of 6 (4+2) and the pairwise combinations of the 4-length and 2-length compositions of 4, respectively: {1+1+1+1} and {1+3, 3+1, 2+2}.

As another concrete example, let us also briefly consider the setting shown in (5d), where the 6 syllables of [a'wɪ.ɬɔ.wa 'rə.ŋə] are distributed across three measures with two syllables in each measure. Given that there exists three 2-length compositions of 4 ({1+3, 3+1, 2+2}), there are  $3^3 = 27$  possible settings with two syllables in each of the three measures (for brevity only one is shown here). Of relevance here is that the settings in (5a–c) and (5d) differ in the total number of measures across which the syllables are distributed (2 vs. 3, respectively). As *Akwelye* has a variable-length template, a constraint-based analysis must consider settings of various lengths for each song line. The use of integer compositions allow us to generate settings with a variable number of total measures in a principled manner.

This method of candidate generation permits us to compute all possible settings for song lines without modifications to the syllables, i.e. the faithful sung lines in *Akwelye*. For computing all possible settings for both faithful and non-faithful options, we simply add a pre-processing stage to the generation of candidate settings via integer compositions. As shown below in (6), a given input text [a'wɪ.ɬɔ.wa 'rə.ŋə] is modified in all 3 of the attested modifications and a copy of the original input is also retained (i.e. as the faithful option). Candidate settings are then exhaustively generated for each of these 4 options and the resulting sets are combined into a single candidate set for that given input. This process was carried out for all 89 song texts, producing a list of 607,543 candidate settings in total.

(6) Faithful and non-faithful ways of setting an *Akwelye* song text.



**3.3 Data and methods** The analysis sketched out in the next section is able to correctly predict the attested winner for 71 of the 89 song texts (79.7% coverage). The constraint set and ranking were verified in OTSoft (Hayes et al., 2013). The data and code used to generate the OTSoft input tableau along with the OTSoft output have been made available on a Zenodo archive (<https://zenodo.org/record/4421813>). Additionally, to enable reproducibility of the current analyses as well as for use by others, the candidate generation process has been made available as an R package: <https://github.com/fauxneticien/akwelye>. Thus, only a brief methodological summary is provided here.

First, *Akwelye* song lines were converted into a working representation as a sequence of syllables, e.g. (6a) 'wɪ.ɬɔ.wa#rə.ŋə' → '#σσσ#σσ', adapting a notation from Hayes (2009). Second, for each song text, candidates were exhaustively generated using the method described in the previous subsection. An example candidate for the input '#σσσ#σσ' is a string value of the form '#σσσ.#σσ..|', which represents the same setting as shown in (6a-i). Here, '|' represents measure boundaries, '#' prosodic boundaries, and '.' additional grid positions associated with the previous syllable. Thus, '#σσ..|' in the second measure represents a setting that sets the line-final syllable [ŋə] to three grid positions in total. Third, the candidate set was ran through the constraint set which results in a tabular output compatible with standard OT software (e.g. OTSoft: Hayes et al., 2013). The constraint set is implemented as a series of regular expressions which count the number of times an illicit expression occurs within a candidate string.

Finally, constraint ranking was verified in OTSoft using the process of Constraint Rank Demotion. To decrease computation time in OTSoft, we used a subset of the data containing only candidates with unique violation profiles, using a similar culling procedure for sets of tied candidates as employed in Hayes (2009: fn.

8). Note that only various clusters of losing candidates had identical violations. That is, while the constraint set was not able to differentiate between all possible candidates from a pool of 607,543, it was nevertheless able to select a unique winner for every one of the 89 inputs — and the correct winner for 71 of them.

### 4 Optimality Theoretic analysis

Beginning with the lowest-ranked constraint, FINAL IS LONG accounts for the fact that in *Akwelye* measure-final syllables are set to multiple grid positions, while all other syllables are set to one grid position. In this way, length is used to indicate the measure boundary, in a way that mirrors English text-setting where the strong position is lengthened (Hayes, 2009). Mirroring Hayes’ (2009) STRONG IS LONG constraint, violations for FINAL IS LONG is assigned by subtracting from 3 the number of positions to which each measure-final syllable is set (e.g. as [gi] in (9a) is set to only 2 positions, there is a violation of 1 = 3 - 2). In addition to FINAL IS LONG, the SALIENT constraint targets line-final syllables to ensure they are set to the longest duration within the line (where possible). Relatedly, \*FILL ensures that each measure contains at least one syllable set to multiple grid positions, prohibiting settings such as (10a), where each of the four grid positions are occupied by a different syllable.

- (7) Metrical constraints for *Akwelye* text-setting:
  - a. \*INCREASE: number of syllables set to consecutive measures must not increase.
  - b. \*STRESSLESS IN STRONG: no unstressed syllables in metrically strong positions.
  - c. \*FILL: measures must contain empty positions.<sup>1</sup>
  - d. SALIENT: last measure must have fewest number of syllables.
  - e. FINAL IS LONG: every measure-final syllable must be long (= set to 3 grid positions, W M W).
- (8) Phonological constraints for *Akwelye* text-setting:
  - a. MAX(V): all input vowels appear in output.
  - b. LINEARITY(ROOT): linear order of syllables within roots from input is retained in output.
- (9) Illustrations for \*STRESSLESS IN STRONG, LINEARITY(ROOT), MAX(V), SALIENT, and FINAL IS LONG.

		*INCREASE	*STRESSLESS IN STRONG	LINEARITY(ROOT)	*FILL	MAX(V)	SALIENT	FINAL IS LONG
/a[an]kwə əɾəpə arə+ŋə/ bloodwood_tree blossom put+RPST [a.'an.gi'ri.na'rə.ŋə]								
a. FAITH	S W M W S W M W S W M W [ a [an] gi [ri] na [rə] ŋə ]		*!				*	*
b. C-EPN	S W M W S W M W S W M W [ la [an] gi [ri] na [rə] ŋə ]						*	*
c. S-TRN	S W M W S W M W S W M W [ rə [an] gi [ri] na [rə] ŋə ]			*!				*
d. v-DEL	S W M W S W M W S W M W [ [an] gi [ri] na [rə] ŋə ]					*!	*	

- (10) Illustrations for \*INCREASE and \*FILL.

/kwəɾəjmpə-la-cə arə+ŋə/ ancestral_woman-ERG-1sg.ERG put+IMP [ku.ri.mpə.la.ca.'rə.ŋə]							
a. FAITH	S W M W S W M W [ ku rim pə la ca 'rə ŋə ]		*		*!		***
b. FAITH	S W M W S W M W S W M W [ ku rim pə la ca [rə] ŋə ]		*			*	*
c. FAITH	S W M W S W M W S W M W [ ku rim pə la ca [rə] ŋə ]	*!	*			*	*

<sup>1</sup> Note here, by ‘empty’ we mean a position that does not have a note attack; instead the previous syllable is held for the duration of two metrical positions.

The standard phonological constraint MAX(V) requires that vowels in the input be realised in the setting. Thus, (9d), which elides the line-initial /a/ violates MAX(V). By contrast, while (9b) and (9c) are consonant-initial lines, notice that these lines do not violate MAX(V) as the vowel is not deleted, but realised as the second segment in the line (e.g. /a/ → [la]).

In the current analysis, the three undominated constraints are \*INCREASE, \*STRESSLESS IN STRONG, and LINEARITY(ROOT). \*INCREASE prohibits settings where the number of syllables in a later measure is greater than the number of syllables set in an earlier measure, as illustrated by (10c), where the two syllables are set to the first measure and three in the second measure.

Recall that syllable transfer is a process where the line-final syllable from the input is realised in the line-initial position in the output setting. As illustrated by (9c), the line-final /ŋə/ from the input is transferred to the line-initial position, where it is realised as [ŋa] (/ŋə/ + /a/ forms [ŋa] via standard vowel hiatus resolution). Syllable transfer is not permissible for this song text as the utterance-final syllable /+ŋə/ +IMP is a root-level suffix. LINEARITY(ROOT) requires that the linear order of syllables within a root in the input is also maintained in the output. In the case of (9), /arə+ŋə/ being realised as [ŋə...arə] violates this constraint. Crucially, for this constraint to function, it must be able to differentiate between word-level suffixes such as /lə/ and root-level suffixes such as /+ŋə/. That is, the input to text-setting process cannot simply be a surface form such as [a'lan.gi.ri.na'rə.ŋə] without any morphological information.

Finally, \*STRESSLESS IN STRONG (Hayes, 2009) prohibits the strong position from being occupied by a stressless syllable. As illustrated in (9a), the faithful candidate which retains the line-initial /a/ results in a setting that places an unstressed [a] in the line-initial strong position. As such, (9a) is assigned a violation by \*STRESSLESS IN STRONG and is eliminated by the constraint. All other candidates do not violate this constraint, as their line-initial syllable is stressable, either through omitting the line-initial /a/ from the input, as in (9d), or placing a consonant in front of this vowel, e.g. via consonantal epenthesis as in (9b).

The ranking illustrated in the tables above was discovered through the Constraint Rank Demotion process in OTSoft. Here we provide an overview of the key constraint interactions that give rise to the ranking. There are attested lines in *Akwelye* which violate \*FILL by leaving no empty grid positions, as shown below in the first measure of (11b). Thus, \*FILL must be ranked below \*STRESSLESS IN STRONG, which prefers (11b) over (11a), which maintains an empty grid position in each measure at the cost of misaligning stressed syllables and strong beats (e.g. setting unstressed [pə] in the strong position of the second measure).

(11) Ranking illustration: \*STRESSLESS IN STRONG » \*FILL

35: aɪ'lə.ŋən.tar.pə'ŋu.rə.lə'wi.jɔw												*STRESSLESS IN S	*FILL	
a. V-DEL	S	W	M	W	S	W	M	W	S	W	M	W	! **	
	lə	ŋən	tar		pə	ŋu	rə		lə	wi	jɔw			
b. V-DEL	S	W	M	W	S	W	M	W	S	W	M	W		*
	lə	ŋən	tar	pə	ŋu	rə	lə		wi	jɔw				

As shown below in (12), both SALIENT and MAX(V) prefer the losing candidate (12b), which leaves no empty grid position in the first measure. This is due to the fact that the winning candidate (12a) omits the line-initial vowel [e], thus violating MAX(V). Additionally, as both measures contain three syllables, the same candidate violates SALIENT, which requires the final measure to contain strictly fewer syllables than those preceding it. Thus, to select the correct winner, SALIENT and MAX(V) must be ranked below \*FILL.

(12) Ranking illustration: \*FILL » MAX(V), SALIENT

3: e'lə.tɔ.wə'lə.tɔ.wə							*FILL	MAX(V)	SALIENT		
a. V-DEL	S	W	M	W	S	W	M	W		*	*
	lə	tɔ	wə		lə	tɔ	wə				
b. C-EPN	S	W	M	W	S	W	M	W	!*		
	le	lə	tɔ	wə	lə	tɔ	wə				

As shown below in (13), FINAL IS LONG prefers the losing candidate (13a), which omits the initial vowel [a], allowing both measure-final syllables to be set to 3 grid positions. Omitting the vowel, however, violates MAX(V), and having a setting with equal numbers of syllables in both measures violate SALIENT. Thus, to select the correct winner, FINAL IS LONG must be ranked below SALIENT and MAX(V).

## (13) Ranking illustration: MAX(V), SALIENT » FINAL IS LONG

14: a'ra.ca 'ra.cə		MAX(V)	SALIENT	FINAL IS LONG																
a. V-DEL	<table border="1"> <thead> <tr> <th>S</th> <th>W</th> <th>M</th> <th>W</th> <th>S</th> <th>W</th> <th>M</th> <th>W</th> </tr> </thead> <tbody> <tr> <td>ra</td> <td>ca</td> <td></td> <td></td> <td>ra</td> <td>cə</td> <td></td> <td></td> </tr> </tbody> </table>	S	W	M	W	S	W	M	W	ra	ca			ra	cə			*	*	
S	W	M	W	S	W	M	W													
ra	ca			ra	cə															
b. C-EPN	<table border="1"> <thead> <tr> <th>S</th> <th>W</th> <th>M</th> <th>W</th> <th>S</th> <th>W</th> <th>M</th> <th>W</th> </tr> </thead> <tbody> <tr> <td>na</td> <td>ra</td> <td>ca</td> <td></td> <td>ra</td> <td>cə</td> <td></td> <td></td> </tr> </tbody> </table>	S	W	M	W	S	W	M	W	na	ra	ca		ra	cə					*
S	W	M	W	S	W	M	W													
na	ra	ca		ra	cə															

We now turn to the 18 song lines that our current analysis cannot account for. For 5 of the 18 song lines, the attested candidate was realised with vowel deletion rather than consonant epenthesis or syllable transfer. All of these song texts commenced with an initial /auə/ sequence underlyingly. Of relevance here is that this phoneme /u/ is a highly-restricted phoneme that can only occur as the first onset of a word (Koch, 1997; Turpin & Ross, 2012). Thus, non-faithful forms resulting from consonant epenthesis or syllable transfer violate this restriction, e.g. /auə/ → \*[lauə...]. The addition of phonotactic constraints enforcing such restrictions would allow us to account for these song lines.

For another 5 of the 18 song lines, the attested candidate set an unstressed vowel in the strong position. As mentioned above, \*STRESSLESS IN STRONG is an undominated constraint in the current analysis. A consequence of this is that a winning candidate may never place an unstressed syllable in the strong position. Yet, as shown by (14a), the attested setting for certain song lines do place an unstressed syllable in the strong position, e.g. the syllable [cij] in the strong position of the second measure. Such candidates are eliminated by \*STRESSLESS IN STRONG, and another candidate that does not violate the constraint is incorrectly selected as the winner, e.g. the transfer candidate in (14b). As our analysis only considers line-level constraints, there may be higher, verse-level considerations which overrule constraints such as \*STRESSLESS IN STRONG, thus accounting for attested settings such as (14a). Indeed, one source of evidence for verse-level constraints operating in *Akwelye* is that for one of the five lines the sung form has two different settings where its first iteration is different to its repetition in the verse in the same verse. This process is nicely illustrated by Ken Hale in relation to a neighbouring Anmatyerr song (Hale, 1984:261–262).

## (14) Erroneous prediction resulting from an undominated \*STRESSLESS IN STRONG constraint.

56: a.'a.cij 'la.cij-la-rə		*STRESSLESS IN STRONG	...																
⊗ a. V-DEL	<table border="1"> <thead> <tr> <th>S</th> <th>W</th> <th>M</th> <th>W</th> <th>S</th> <th>W</th> <th>M</th> <th>W</th> </tr> </thead> <tbody> <tr> <td>la</td> <td>cij</td> <td>la</td> <td></td> <td>cij</td> <td>la</td> <td>rə</td> <td></td> </tr> </tbody> </table>	S	W	M	W	S	W	M	W	la	cij	la		cij	la	rə		!*	
S	W	M	W	S	W	M	W												
la	cij	la		cij	la	rə													
⊙ b. S-TRN	<table border="1"> <thead> <tr> <th>S</th> <th>W</th> <th>M</th> <th>W</th> <th>S</th> <th>W</th> <th>M</th> <th>W</th> </tr> </thead> <tbody> <tr> <td>ra</td> <td>la</td> <td>cij</td> <td></td> <td>la</td> <td>cij</td> <td>la</td> <td></td> </tr> </tbody> </table>	S	W	M	W	S	W	M	W	ra	la	cij		la	cij	la			
S	W	M	W	S	W	M	W												
ra	la	cij		la	cij	la													

For another 5 of the 18 song lines, the attested candidate was found to be the second-best candidate, and this fact may be explainable by a probabilistic grammar that is not implemented by the ‘Classical OT’ approach we have adopted here. For example, using a stochastic maximum entropy (maxent) grammar, Hayes & Moore-Cantwell (2011) examined how well their grammar could predict the correct scansion for lines in a corpus of Gerard Manley Hopkins’ sprung rhythm poems. One finding from this analysis was how on occasion top two candidate scansions were assigned considerably higher probabilities (e.g. 1st: 0.665, 2nd: 0.332 Hayes & Moore-Cantwell, 2011, 286) than the subsequent candidate scansions (e.g. 3rd: 0.002, 4th: 0.001). As non-faithful options in *Akwelye* do not occur with equal frequency, a maxent grammar may yield better coverage of the *Akwelye* data than our current approach. It remains to be seen, however, whether exhaustive candidate generation interferes with a maxent grammar, given all candidates are assigned some amount of probability.

Finally, for 3 lines, there does not appear to be any obvious pattern to their exceptional nature. For example, for one song line the attested setting deletes both the word-initial vowel [a] and the word-final syllable (unlike any other song line). For another two lines, a line-initial syllable is epenthesised to the sung form, despite the underlying form being a consonant-initial line. Having 3 of 89 (3% of song lines) apparently true exceptions, however, is plausible given the artistic nature of verbal art.

## 5 Conclusion

This paper brings to light a cross-linguistically unusual set of text-setting strategies and shows that the song-specific selection of a strategy can be accounted for through a relatively generic set of text-setting constraints, a number of which were adapted from or had already been proposed by previous research (Hayes, 2009; Kiparsky, 2006). As a descriptive contribution, this paper has shown how the metrical template of *Akwelye* signals measure boundaries through increases in the rhythmic duration of the measure-final syllables. As a methodological contribution, this paper developed a method of exhaustive generation of text-setting candidates and has made the code used in this generation process freely available for use in the analysis of other song traditions. As a theoretical contribution, this paper puts forward an analysis of the prosodic, morphological, and rhythmic contexts that permit the phenomenon of syllable transfer in *Akwelye*. Additionally, this paper has highlighted how further investigation of text-setting in *Akwelye* may benefit from and thus also provide insights to the broader investigation into the complex interaction between language, meter, and music.

## References

- Baker, Brett (2014). Word structure in Australian languages. Koch, Harold & Rachel Nordlinger (eds.), *The Languages and Linguistics of Australia: A comprehensive guide*, Walter de Gruyter, Berlin, Boston: De Gruyter Mouton, 139–213.
- Baker, Brett & Mark Harvey (2003). Word structure in Australian languages. *Australian Journal of Linguistics* 23:1, 3–33.
- Clunies Ross, Margaret, Tamsin Donaldson & Stephen Wild (1987). *Songs of Aboriginal Australia*. Oceania Publications, Sydney.
- Dell, François & John Halle (2009). Textsetting as constraint conflict. Aroui, Jean-Louis & Andy Arleo (eds.), *Comparing Musical Textsetting in French and in English Songs*, John Benjamins, Amsterdam, vol. 2 of *Internal and External Variation in Linguistics*, 63–78.
- Hale, Ken (1984). Remarks on creativity in Aboriginal verse. Kassler, J. C. & J. Stubington (eds.), *Problems and solutions: Occasional essays in musicology presented to Alice M. Moyle*, Hale and Iremonger, Sydney, 254–262.
- Halle, John & Fred Lerdahl (1993). A generative textsetting model. *Current Musicology* 55, 3–23.
- Hayes, Bruce (2009). Textsetting as constraint conflict. Aroui, Jean-Louis & Andy Arleo (eds.), *Towards a Typology of Poetic Forms: From Language to Metrics and Beyond*, John Benjamins, Amsterdam, vol. 2 of *Internal and External Variation in Linguistics*, 43–62.
- Hayes, Bruce & Claire Moore-Cantwell (2011). Gerard Manley Hopkins' sprung rhythm: corpus study and stochastic grammar. *Phonology* 28:2, 235–282.
- Hayes, Bruce, Bruce Tesar & Kie Zuraw (2013). OTSoft 2.5. Software package, URL <http://www.linguistics.ucla.edu/people/hayes/otsoft/>.
- Kiparsky, Paul (2006). A modular metrics for folk verse. Drescher, B. E. & N. Friedberg (eds.), *Formal approaches to poetry*, Mouton de Gruyter, vol. 11, 7–49.
- Kiparsky, Paul (2020). Stress, meter, and text-setting. Gussenhoven, C & A Chen (eds.), *The Oxford Handbook of Language Prosody*, Oxford University Press, Oxford, 657–675, URL [https://web.stanford.edu/~kiparsky/Papers/handbook\\_2016.pdf](https://web.stanford.edu/~kiparsky/Papers/handbook_2016.pdf).
- Koch, Harold (1984). The category of 'associated motion' in Kaytej. *Language in central Australia* 1:1, 23–34.
- Koch, Harold (1997). Pama-Nyungan reflexes in the Arandic languages. Tryon, D & Michael Walsh (eds.), *Boundary rider: Essays in honour of Geoffrey O'Grady*, Research School of Pacific and Asian Studies, Australian National University, 271–302.
- Koch, Harold (2004). The Arandic subgroup of Australian languages. Bowerman, Claire & Harold Koch (eds.), *Australian languages: Classification and the Comparative Method*, John Benjamins Publishing Company, Amsterdam, 127–150.
- Lieberman, Mark & Alan Prince (1977). On stress and linguistics rhythm. *Linguistic Inquiry* 8:2, 249–336.
- Sutton, Peter (1987). Mystery and change. Clunies Ross, M & T Donaldson (eds.), *Songs of Aboriginal Australia*, University of Sydney, 77–96.
- Turpin, Myfany (2005). *Form and meaning of Akwelye: a Kaytetye women's song series from Central Australia*. Ph.D. thesis, University of Sydney.
- Turpin, Myfany (2007a). Artfully hidden: text and rhythm in a Central Australian Aboriginal song series. *Musicology Australia* 29:1, 93–108.
- Turpin, Myfany (2007b). The poetics of Central Australian song. *Australian Aboriginal Studies* :2, 100–115.
- Turpin, Myfany & Alison Ross (2004). *Awelye Akwelye: Kaytetye women's songs from Arnerre*, Central Australia. Audio CD & Cassette. Papulu Apparr-kari Language and Culture Centre, Tennant Creek. Recordings by Grace Koch, Linda Barwick and Myfany Turpin, commentary by Myfany Turpin and Alison Ross.
- Turpin, Myfany & Alison Ross (2012). *Kaytetye-to-English dictionary*. IAD Press, Alice Springs, N.T., Australia.