Deletion and Insertion: Two pathways towards Optimality

MARINA TZAKOSTA
Leiden Center of Linguistics/HIL

1. INTRODUCTION
In the present paper I will discuss two phonological processes, which are very common in child language, segmental deletion and insertion, in order to investigate the degree to which they facilitate the process of language acquisition. Deletion has been focused on the study of cluster and final consonant reduction as well as syllable reduction, known as truncation. Insertion, syllabic and segmental, is a less well-studied process, since it is less common than deletion in child language. In this study I will discuss deletion and insertion as segmental processes within the domain of the syllable. I will briefly refer to them as processes taking place when they are directly affected by the stress system of the language acquired. The data have shown that they are systematic processes and not random errors. They are part of the process of speech simplification (Oller 1974) so that the children can form unmarked structures (McCarthy and Prince 1994).

I will use the data of spontaneous speech from three children Dionisis (D), Marilia (M) and Bebis (B), who acquire Greek as their native language, varying in age from 1;10 to 3;05,23. The data come from a longitudinal corpus of data from 10 children collected for more than a year in Rethymno (Crete). The data were collected in weekly intervals and they are transcribed in the International Phonetic Alphabet (IPA).

The paper is structured as follows: in section 2 I will present some cross-linguistic research findings, in section 3 I will present some findings from the acquisition of Greek and I will describe the emergence of these two processes in detail. I will examine the quality of deleted or inserted segments with respect to place and manner of articulation and I will discuss their intra and inter characteristics, i.e. the relations that might hold within the same process as well as those that hold among different processes. I will also discuss the position of the syllables, which favor the emergence of deletion and insertion as well as the quality of these syllables (e.g. stressed or unstressed, initial or final syllables). In section 4 I will adopt Optimality Theory, in order to account for our data and I will demonstrate the reasons why I think that OT is the best model to account for language acquisition. Finally, I will present our conclusions in section 5.
2. CROSS-LINGUISTIC RESEARCH FINDINGS
As in child language, Creole languages tend to favor simple CV syllables (Alber and Plag 1999). Evidence for the same strategies is also very common in loanwords where vowel epenthesis is triggered and affected by the stress system of the borrowing language (Broselow 1999), as well as dialects (Miglio 1998).

In the literature, two kind of insertion have been distinguished: (a) anaptyxis, i.e. the development of a vowel in O(bstruent)+S(onorant) clusters and (b) prothesis, i.e. the emergence of a vowel before S(ibilant)+S(top) clusters (Fleischhacker 2000).


Cluster reduction is considered to be more common in onset position than in coda position (Gnanadesikan 1995, 1996). Barlow (1997) and Lukaszewicz (2000), who analyze reduction and assimilation processes in English and English and Polish respectively, indicate that the selection of the surviving consonant in cluster simplification does not depend on the position of that consonant within the onset, does not affect the segments that are not yet acquired and is not restricted to a particular class of sounds. On the contrary, they claim that it is based on the relative sonority of the segments in the target onset, i.e. the less sonorous segment is retained, while the more sonorous segment is deleted.

With respect to final consonant deletion Vihman (1978) reports the special behavior of glides. When a child substitutes a glide for a consonant the change is treated as deletion. She does not explain why, though. Furthermore, theoretically speaking, Goad (1998), Goad and Brannen (to appear), and Rose (2000) assume that final consonants are considered to be heads of empty headed syllables, position which is in accordance with the need for CV syllables. According to Stemberger (1996) insertion mainly takes the shape of vowel insertion, pursuing the CV syllabic shape. The data from Greek demonstrate the emergence of consonant insertion for the same reason.

3. EVIDENCE FROM GREEK
Deletion, and insertion are common processes in the speech of children who acquire Greek too. Still, deletion is the process, which has mainly received most attention (Kappa 1997, Tzakosta 1999, 2000). In my study I will address the following issues: Which process is more common and more frequent? What is the nature of the deleted or inserted segment? Do we find
copy or default insertion? What is the favored position of the deleted or inserted segments (word initially, internally or finally)? To what degree are insertion and deletion triggered by the stress system of Modern Greek?

**CLUSTER REDUCTION**

(1) BEBIS CORPUS

(a) /kou`koutsi/ → ['kuti] «pit»
(b) /spiti/ → ['piti] «house»
(c) /fruta/ → ['futa] «fruits» (1;10)
(d) /pe`xmiðja/ → ['duja] «toys»
(e) /a`ftaci/ → ['dida] «ears (dim.)» (1;11,07)
(f) /aspro/ → ['apro] «white»
(g) /vlepi/ → ['lepi] «he sees» (1;11,29)
(h) /sti`lo/ → [to`to] «pen (accus.)» (2;01,05)

(2) DIONISIS CORPUS

(a) /gyrafo/ → ['pao] «I’m writing» (2;1)
(b) /ekatsa/ → ['ekata] «I sat down» (2;01,16)
(c) /taçi`ðromos/ → [taç`yomoθ] «postman» (2;2,12)
(d) /stratio`taci/ → [taçi`taci] «soldier (dim.)» (2;03,07)
(e) /ðromos/ → ['ðomoθ] «road» (2;04,17)
(f) /kal`tula/ → [ka`tula] «sock (dim.)» (2;05,08)
(g) /nostimo/ → ['noθimo] «delicious» (2;06,08)
(h) /xiti`pame/ → [ti`bame] «we break sth.» (2;07,06)

(3) MARILIA CORPUS

(a) /a`fti/ → [va`ti] «she»
(b) /sxo`lio/ → [xo`lio] «school»
(c) /bro`sta/ → [bo`ta] «in front of» (2;07,06)
(d) /pa`putsja/ → [pa`puca] «shoes» (2;07,22)
(e) /tsixla/ → ['txlxa] «chewing gum» (2;08,15)
(f) /trome/ → ['tome] «we eat» (2;11,18)
(g) /trayu`ðane/ → [ta`yane] «they sing»
(h) /espase/ → ['epe se] «it broke» (3;05,23)

**FINAL CONSONANT DELETION**

(4) BEBIS CORPUS

(a) /sciolo/ → ['cilo] «dog»
(b) /ðosto/ → ['ðoto] «give it»
(c) /kokoras/ → ['kokora] «rooster» (1;11,07)
(d) /ðen a`niji/ → [ne`niji] and [en a`niji] «it doesn’t open» (1;11,15)
(e) /valto/ → ['vato] «put it»
(f) /kar`puzi/ → ['kuzi] «watermelon»
(g) /ar`kuDia/ → [a`kuDia] «bear» (1;11,28)
(h) /kurjes/ → ['kurje] «swings»
(5) DIONISIS CORPUS
(a) /πιρυοσ/ → [ˈpiyro] «tower» (1;11,29)
(b) /πορτο/ → [ˈpato] «take it» (2;01,05)

(6) MARILIA CORPUS
(a) /οργακι/ → [joˈgaki] «George (dim.)» (2;07,06)
(b) /εροδε/ → [ˈerode] «they come» (2;07,22)
(c) /ο παπος/ → [o paˈpu] «grandfather»
(d) /ο διμιτηρ/ → [o diˈmiti] «proper name»
(e) /πορτα/ → [ˈpata] and [baˈbata] «take it» (2;08,07)
(f) /κανυν/ → [ˈkanun] «they do»

(7) BEBIS CORPUS
(a) /επεσε/ → [ˈpesa] «it fell down»

(8) DIONISIS CORPUS
(a) /δεν ʹεξι/ → [ˈe lεxi] «he doesn’t have»
(b) /δεν ʹεξω/ → [ˈe lεxɔ] «I don’t have» (2;01,23)
(c) /δεν άφινι/ → [ˈte fini] «s/he doesn’t leave»

(9) MARILIA CORPUS
(a) /αφι ʹειν/ → [ˈtine] «it is her» (2;07,22)
SEGMENT INSERTION
(10) BEBIS CORPUS
(a) /'etsi/ → ['neti] «like that (adv.)»
(b) /'aloyo/ → ['nayoyo] «horse»
(c) /a'fto/ → [cpi'to] «this» (1:11,28)
(d) /'natin/ → ['natine] «here she is»
(e) /u'ra/ → ['dula] «tail» (2:01,05)
(11) DIONISIS CORPUS
(a) /'den 'exo/ → ['ve 'lexo] «I don’t have» (2:01,23)
(b) /'istera/ → ['titela] «later (adv.)» (2;2,12)
(c) /aste'racal/ → [yate'laa] «stars (dim.)» (2;04,17)
(d) /'partin/ → ['patine] «take her»
(e) /vi'via/ → [volo'via] «books» (2:07,06)
(f) /e'kxatmis/ → [ce'kxatmis] «evaporation» (2;09)
(12) MARILIA CORPUS
(a) /a'fti/ → ['vati] «she» (2:07,06)
(b) /aero'plano/ → [vao'pano] «aeroplane» (2;07,22)
(c) /a'fto/ → [te'to] «this one» (2:08,07)
(d) /e'go/ → [go'go] «I» (2;08,15)
(e) /ja'urtca/ → [ja'ute-la] «yogurt»
(f) /'aloyo/ → ['dilo jo] «horse» (2;08,22)

The above data lead to the following conclusions:
1. Deletion in any of its shapes is a more common process than insertion as is shown in (13).¹

(13)

<table>
<thead>
<tr>
<th></th>
<th>Deletion (in general)</th>
<th>Insertion (in general)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bebis</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>Dionisis</td>
<td>25</td>
<td>6</td>
</tr>
<tr>
<td>Marilia</td>
<td>28</td>
<td>6</td>
</tr>
</tbody>
</table>

(13) demonstrates that even Marilia, who is the older child, finds it easier to delete rather than insert segments. This might be due to the fact that children prefer to truncate their forms rather than lengthen them. Children insert segments mainly when the prosodic shape of their words does not change, e.g. when the number of the syllable remains the same compared to that of the adult form ((10a), (10b), (11b), (12d)), or to a lesser degree, when it is necessary for the children to do so in order to form CV syllables ((11d), (11e), (12e)).

¹ The numbers in the tableau indicate the cases found in the data not percentages.
2. Generally speaking, Greek has the priority to simplify complex structures. More frequent cluster reduction, compared to coda deletion, might also be due to the fact that coda consonants, i.e. closed syllables are very limited in Greek. Syllables, which are closed with more than one segment, do not emerge in Greek (Kappa 1995). At the stage we are studying, most clusters are simplified to the unmarked stop member of the cluster (e.g. (1b), (2d), (3c), (3g)), or the less marked fricative, e.g. labials are preferred over liquids (e.g. (1c)), or to a stop which substitutes a fricative (e.g. (2c)). Clusters consisted of 3 members are simplified to 2 consonants with [s] being more vulnerable to deletion, since it is considered to be extrasyllabic (e.g. (1f)). Syllabic codas are also deleted in all positions, either word medially or finally. [n],[s],[r],[l] are the segments which can occupy the coda position in Greek (e.g. (4a), (4e), (4f), (4i), (5a), (5d), (5e), (5l), (6b), (6c), (6j)). As simple segment deletion I call those deletions, which do not fall within cluster reductions and final consonant deletions, such as single onset deletion (e.g. (8a), (8c), (9a)) or deletion of syllables, which consist of only one vowel (e.g. (7a), (7b), (8c)). Segment insertion takes the shape of onset insertion in the beginning of the word (such as (10a), (10b), (10e), (11b), (12d)), vowels breaking clusters in all positions within the word (e.g. (10c), (11e), (12e)), vowels following codas in order to form CV syllables but much rarer (e.g. (10d)). The inserted consonants are stops mainly labials and coronals. Cross-linguistically, Lombardi (1998) reports that laryngeals are as common as coronals in insertion processes. Smolensky (1993), also notices that ‘Laryngeals vs. Coronals are in a fight to death’. Grijzenhout and Joppen (2000) report that in German initial affricates are simplified, initial fricatives are deleted or realized as plosives. They furthermore claim that C-deletion in CV sequences are very common, since V syllables are as favorable as CV syllables in German child speech (see Velten (1943) for English and Costa and Freitas (1998) for Portuguese).

3. In Greek I have noticed that the inserted consonant can be either default, i.e. an unmarked segment is produced (cf. Kitto and de Lacy 1999) ((10b), (10e), (11c)), or it can be similar to another segment in the word with respect to place and/or manner of articulation (copy insertion) (e.g. (10a), (11b), (12c), (12d)).

4. Deletion, mainly in the shape of cluster reduction, is very common in all positions, word initially, medially, or finally, in stressed or unstressed syllables. In a later stage (e.g. (11f)) clusters start being retained in stressed syllables, though. Stressed syllables tend to be faithful to the input, namely they keep as many of their features and segments as possible.

5. Since Greek is (unlike Ancient Greek) a quantity insensitive language, cluster or coda deletion as well as segment insertion seem not to specially
affect the accentual shape of the early words. The primary aim of deletion and insertion as segmental processes is the accomplishment of CV syllables. Still, examples such as (7a), (7b), (8c) which are cases of syllabic deletion indicate the role these processes can play in the shape that early prosodic words take, i.e. the binary trochaic foot in Greek (Kappa 1998).

4. LANGUAGE ACQUISITION AND OPTIMALITY THEORY
The theoretical framework we adopt in order to account for our data is Optimality Theory (OT, Prince and Smolensky 1993). According to OT, the set of output candidate forms, which are based on an underlying structure, are evaluated on the basis of a set of universal constraints. These are faithfulness constraints, which require faithfulness to the base and well formedness constraints, which favor simple structures. These universal constraints are ranked in a hierarchy in every language, which constitutes the language’s grammar.

Every language’s hierarchy is reflected on the adult speech. Adult language constitutes the input for children acquiring their native language. Still, child language is hardly faithful to the input; what the child cares about is simple and unmarked forms. As a result, child language differs from adult language as far as the constraint ranking is concerned. And that’s why well formedness constraints are ranked higher, while faithfulness constraints are ranked lower in the hierarchy of child speech.

But then how does the child acquire and adopt the adult language hierarchy, which is the hierarchy of his native language? The problem is settled by means of constraint demotion. During the process of language acquisition the child starts acquiring marked and complex structures. That leads to the re-ranking of the constraint, with faithfulness constraints being ranked higher over time. This also gives an answer to the problem posed by McMahon (2000) who questions whether the OT constraint system is regarded as fixed or flexible. We believe that language acquisition could be used as an example of ‘language evolution’. Constraint re-rankings, which lead from unmarked to marked and ‘difficult’ structures show the different stages that the acquisition of a language comes through. So, the OT system cannot be considered anything but flexible.

With respect to our topic, the constraints I adopt are the following:

**FAITHFULNESS CONSTRAINTS**

MAX: segments of the input must be preserved in the output (no deletion)

(PARSE)
DEP: segments of the output must be dependent on the input (no insertion) (FILL)\(^2\)

**WELL-FORMEDNESS CONSTRAINTS**
*COMPLEX (O): no complex onsets
*COMPLEX (Rhyme): no branching is allowed in the rhyme\(^3\)
CV: syllables must be CV
[-CONT]: segments are [-CONT]

I have already mentioned that deletion, in any of its forms form, i.e. syllabic or segmental, is more frequent than insertion. I assume that this is reflected in the following ranking

cluster deletion >> coda deletion >> onset insertion >> vowel insertion

This order leads me to conclude that markedness constraints are higher ranked than faithfulness constraints. So, constraints such as *COMPLEX preventing marked forms will be highly ranked. If I assume that the inputs are candidate outputs too the reasons they are not selected are because (a) well formedness constraints are higher ranked than faithfulness constraints and (b) even when faithfulness constraints, such as MAX and DEP, come into play, they are ranked relatively low so that they allow insertion and deletion. The ranking I propose in order to account for our data is

*COMPLEX (O), *COMPLEX (rhyme) >> CV\(^4\) >> DEP >> MAX >> [-CONT]

According to the proposed hierarchy what is expected is simple forms to emerge in child speech. The constraints *COMPLEX (O) and *COMPLEX (R) are indeterminately ranked, since I consider them to be equally important for the topic I investigate.

---

\(^2\) The difference between MAX and PARSE and DEP and FILL respectively, is that PARSE and FILL are introduced by the Standard OT Theory (Prince and Smolensky 1993), while MAX and DEP are introduced by Correspondence Theory (McCarthy and Prince 1995).

\(^3\) I find it essential to divide the general *COMPLEX constraint in the *COMPLEX (O) and *COMPLEX (Rhyme) constraints, because they are crucial for the selection of the optimal candidate, e.g. either when candidate forms violate one or both of them, or when multiple violation occurs with respect to the forms' complexity, which fatally leads to the rejection of certain candidates.

\(^4\) The CV constraint stands for the general *COMPLEX constraint.
In the following tableaux I will show the operation of deletion and insertion. In tableau 1 I show the constraint ranking with respect to cluster reduction.

Tableau 1

<table>
<thead>
<tr>
<th>Forms / Constraint</th>
<th>*COMPLEX (R)</th>
<th>*COMPLEX (O)</th>
<th>CV</th>
<th>DEP</th>
<th>MAX</th>
<th>[-CONT]</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>[spiti]</code></td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[piti]</code></td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[vlepi]</code></td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[lepi]</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[kal’tsula]</code></td>
<td>*</td>
<td>*</td>
<td>**</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[ka’tula]</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[bro’sta]</code></td>
<td>**</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[bo’ta]</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

The fourth and fifth pair of candidate forms demonstrate that both the highly ranked markedness constraints are violated by the complex candidates, which are not selected as optimal in the end. In the first case, [kal’tsula] violates both *COMPLEX (R) and *COMPLEX (O) and consequently the form is considered to be the optimal output. Also, the form [bro’sta] causes double violation of the *COMPLEX (O) and is excluded from being selected. As a result, the simplest forms [ka’tula] and [bo’ta] are selected respectively.

In tableau 2 I show how the constraint ranking accounts for final consonant deletion.

Tableau 2

<table>
<thead>
<tr>
<th>Forms/Constraints</th>
<th>*COMPLEX (R)</th>
<th>*COMPLEX (O)</th>
<th>CV</th>
<th>DEP</th>
<th>MAX</th>
<th>[-CONT]</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>[scilos]</code></td>
<td>!</td>
<td>!</td>
<td>**</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[cilo]</code></td>
<td></td>
<td></td>
<td>!</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[portoka’li]</code></td>
<td>!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[f’otoka’li]</code></td>
<td></td>
<td></td>
<td>!</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[vlepun]</code></td>
<td>!</td>
<td>!</td>
<td>**</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[lepu]</code></td>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[tutta]</code></td>
<td>!</td>
<td>!</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td><code>[tuta]</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>
The second pair of candidate forms of tableau 2 shows that even though the ranking of *COMPLEX (O) and *COMPLEX (R) is thought to be indeterminate, the optimal output indicates that MG requires the rhyme not to branch, even if the syllabic onset is complex ([portoka`li] vs. [f̩otoka'li]). It is preferred that the onset is complex rather than the coda (cf. more examples in our data).

In tableau 3 I show the constraint ranking with respect to segmental insertion.

<table>
<thead>
<tr>
<th>Forms</th>
<th>*COMPLEX (R)</th>
<th>*COMPLEX (O)</th>
<th>CV</th>
<th>DEP</th>
<th>MAX</th>
<th>[-CONT]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[aste`rac]</td>
<td>*!</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[yate`laca]</td>
<td>*!</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[aloyo]</td>
<td>*!</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[siloyo]</td>
<td>*!</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[vi`via]</td>
<td>*!</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[vol`via]</td>
<td>*!</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[ja`urtca]</td>
<td>*!</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[ja`utε la]</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

In tableau 3 forms like [yate`laca] and [ja`utε la] are preferred in order to break up clusters, delete coda consonants and form CV syllables even at the expense of the prosodic structure. DEP and MAX are equally violated but for different reasons, namely insertion and deletion both occur for the sake of well formedness, i.e. unmarked structures.

5. CONCLUSIONS
I hope that the study of deletion and insertion has shed light on the acquisition of the syllable in Modern Greek. The findings have led me to the following conclusions.

1. Deletion and insertion are very common child language repair strategies, in order to accomplish universally unmarked CV syllables.

2. Deletion and insertion also lead to segmental unmarkedness, i.e. stopping, with respect to manner of articulation, coronality and labiality with respect to place.

3. Deletion is more frequent than insertion.

4. Segmental deletion and insertion are not influenced by the accentual system of Modern Greek, which is not a quantity sensitive language.
5. OT adequately demonstrates the flexibility children show during the process of language acquisition.

ACKNOWLEDGEMENTS
Special thanks to J. van de Weijer for useful comments on an earlier draft of this paper. Also M. Aronoff, P. Hironymous, P. Kiparsky and A. Ralli for the discussion.

REFERENCES


