Lemmes d'itération

Exercice 1 (Lemmes d'itération)
Soit L un langage reconnaissable sur un alphabet A. Montrer les trois propriétés suivantes :

1. Il existe un entier n tel que pour tout mot w de L tel que $|w| \geq n$, il existe trois mots $u_1,\ u_2,\ u_3$ de A^* tels que $w = u_1u_2u_3$, $u_2 \neq \varepsilon$ et $u_1u_2^*u_3 \subseteq L$.
2. Il existe un entier n tel que pour tout mot w de L qui s'écrit sous la forme $w = w_1w_2w_3$ avec $|w_2| \geq n$, il existe trois mots $u_1,\ u_2,\ u_3$ de A^* tels que $w_2 = u_1u_2u_3$, $u_2 \neq \varepsilon$ et $w_1u_1u_2^*u_3w_3 \subseteq L$.
3. Il existe un entier n tel que pour tout mot w de L, pour toute suite d'entiers
$$0 \leq i_0 < i_1 < \cdots < i_n \leq |w|$$
il existe deux entiers $0 \leq j < k \leq n$ tels que si on écrit $w = u_1u_2u_3$ avec $|u_1| = i_j$ et $|u_1u_2| = i_k$ alors $u_1u_2^*u_3 \subseteq L$.

Exercice 2 (Applications des lemmes d'itération)
En utilisant les lemmes d'itération, montrer que ces langages ne sont pas reconnaissables :
- $L_1 = \{w \in \{a,b\}^* \mid w \text{ est un palindrome}\}$
- $L_2 = \{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$
- $L_3 = \{a^{n^2} \mid n \geq 0\}$
- $L_4 = \{a^nb^p \mid n \neq p\}$
- $L_5 = \{ab^{k_1}ab^{k_2}\dots ab^{k_p}a \mid p \geq 0 \text{ et } \exists i > 0\ k_i \neq i\}$ (Langage de Goldstine)

Soit L un langage reconnaissable, les langages suivants sont-ils toujours reconnaissables ?

1. $\{uu \in \{a,b\}^* \mid u \in L\}$
2. $\{u \in \{a,b\}^* \mid uu \in L\}$
3. $\{u \in \{a,b\}^* \mid \exists v \text{ tel que } |u| = |v| \text{ et } uv \in L\}$
4. $\{u \in \{a,b\}^* \mid aubuubabu \in L\}$ (on pourra utiliser la représentation par monoïde)

Exercice 3 (Non équivalence des lemmes d'itération)
Nous allons montrer que les lemmes d'itération de l'exercice 1 ne sont pas équivalents.

1. Montrer que le langage L_2 vérifie la propriété 1 mais pas la propriété 2.
2. Montrer que, si $A = \{a, b, c, d\}$, le langage

$$L_6 = \{(ab)^n(cd)^n \mid n \geq 0\} \cup A^*(aa + bb + cc + dd + ac)A^*$$

vérifie la propriété 2 mais pas la propriété 3.

Exercice 4 (Non exhaustivité des lemmes d'itération)
On dit qu'un mot contient un carré s'il peut s'écrire sous la forme $uvvu'$ avec $v \neq \varepsilon$. Soit L l'ensemble des mots de $\{a, b, c, d\}^*$ qui s'écrivent sous la forme udv avec $u, v \in \{a, b, c\}^*$ et soit $u \neq v$, soit u ou v contient un carré.

1. Montrer que L vérifie la propriété 3 pour $n = 4$.
2. Montrer que L n'est pas reconnaissable. On pourra admettre l'existence de mots sans carré arbitrairement longs (sur un alphabet à au moins 3 éléments) et appliquer un lemme d'itération au complémentaire de L.

Exercice 5 (Langages 1-LL)
Les langages qui vérifient le lemme 3 sont appelés les langages 1-LL. Montrer que la famille des langages 1-LL est fermée par union, concaténation et étoile. Montrer en outre que cette même famille contient strictement l'ensemble des langages reconnaissables.
Indication : On pourra considérer le langage $L_7 = M \cup K$ où

$$M = \{(ab)^n(cd)^n \mid n \geq 1\} \quad \text{et}$$
$$K = A^*(aa + ac + ad + bb + bd + ca + cb + cc + da + db + dd)A^*.$$
